

Between an Analogue Past and a Digital Future: The Evolving Digital Present

Eleanor Selfridge-Field (Stanford)¹

Academe is in turmoil. The digital revolution confronts us with new challenges every day. Should we publish in print or online? Store our work on personal hard disks, on shared institutional servers, or a proprietary storage “cloud”? Teach in a traditional classroom or in virtual online space? We know the positive aspects of the physical world of Academe, but what of the virtual one? The shelf-life of any digital medium in use today is uncertain.

However bewildering these questions may be for our colleagues in literary and historical studies, they are far more bewildering for musicologists. Yet a backward glance can remind us that our discipline is roughly as old as the recording industry (both were in a fledgling state in the 1880s). The history of the musical recordings may suggest a fuzzy version of the future of our discipline. Methods of recording have changed countless times, and each change has necessitated the migration of media to newer formats. Yet the recording industry is still very much alive and so is musicology.

The discomforts of a virtual future owe principally to the diversity of materials on which our discipline is based. Studies of music rely on access to music of all kinds; to printing technology that supports music as well as text, including diacritical marks and alphabets that may be foreign to one’s own language; to all manner of graphical superimposition (e.g. of Schenkerian graphs, lute tablature, or basso continuo figuration); and, in the most fundamental of ways, access to sounding music. While the apparatus in which our scholarship has been reported over the past century are being relinquished, the realities of e-publishing offer little solace, since (at this writing) most of our materials remain too complex to be fully adapted. As the trappings of more than a century of scholarship are cast adrift, the virtues of the luxuriously long age of print recede in a golden glow behind us.

Despite the poor fit between the print culture to which we are accustomed and the digital one to which we are now subjected, opportunities for a richer kind of scholarship, with more built-in efficiencies, abound at every turn. Nearly everyone with an interest in digital musicology recognizes the need for a comprehensive, constantly updated portal for the discipline so that we can stay abreast of these burgeoning possibilities. This has so far proved an elusive goal, although ViFaMusik, the sponsor of this session, offers one of the best starting points (see <http://www.vifamusik.de/startseite.html>). On the broader plane of the future for which we should prepare, we turn our attention bilaterally to (1) digitized source material and (2) digital ways of using it. We then turn, in conclusion, to the challenges to expertise that our emerging virtual world makes palpable.

¹ The author expresses her sincere thanks to Ilias Chrissochoidis, Judith Haug, and Craig Stuart Sapp for important contributions to this text.

I. Silent Sources

What is unambiguous about the term “digitized” sources is their origins: they have all existed in a physical medium. What is ambiguous is what the term means in any given instance. In recent years it has tended to refer to digital images posted online, often in a PDF format. In this case we might infer that the images are accompanied by textual metadata, since digitized images are not (at least without further processing) searchable.

There are myriad ways in which digitized sources (let us take a leaf of music as a starting point) take on additional value or meaning: (a) the music is encoded and thus fully searchable; (b) the music shown on the leaf can be synchronized with a sound recording or a MIDI file, so we can hear it as we peruse the score; or (c) the handwriting and/or watermark associated with this leaf can be identified. Of course there is a multitude of other uses we may make, but this short list serves merely to explain that it is generally useful to differentiate between sources that are directly searchable and those that are not; between graphics and text; between graphics and sound; and between sound and text.

IA. Graphical access and analysis

IA.1. Direct access to music manuscripts

Within historical musicology the gap between access to published and unpublished sources has been immense. The advent of digitized rare materials seems altogether miraculous to those of us who have labored long in libraries and archives in order to gather information from unique sources for which there is no permissible way to photocopy, transcribe, or investigate a watermark. Research that used to require a passport, an air or train ticket, a hotel reservation, a reader’s card, and a No. 2 pencil can in many cases be done in an armchair. Europe’s musical past can be vividly reconstructed on a computer screen from views that are larger and clearer than the poor-quality reproductions that were appearing in much respected books as recently as three decades ago. As one example, I cite Machaut’s exquisitely presented poetry (521 folios in the source I examined virtually from the online holdings of the Bibliothèque Nationale de France (<http://gallica.bnf.fr>). Memories of lessons I once had to grasp from dull grey print emerge as the illuminations of these royal holdings burst into view. We are quickly reminded of Machaut’s importance as a poet by the king’s seal on the spine of some of these leather-bound sources (e.g., <http://gallica.bnf.fr/ark:/12148/btv1b84490444.r=Machaut.langEN>). It is, however, the elaborate illuminations, the detail of the ornamental initials, and the music itself that makes us respond so immediately to the imagery of the “Jugement du Roi” (a song Machaut wrote for his patron, King John of Bohemia, before 1346).²

How dramatically access to such arresting materials promises to change the dynamics of teaching and research! On the one hand, the good condition in which the document now exists attests to seven-and-a-half centuries of care. Yet confinement has imposed great deprivations on all potential readers for three-quarters of a millennium.

The idea inheres in traditional humanities scholarship that our knowledge base consists of a core set of data which grows broader and deeper with each generation. Direct access to original sources may be thought to reduce the need for commentary, although (to continue with the Machaut example) some of the French spellings, plus the details of the illuminations,

² Quoted here from F-Pn, Guillaume de Machaut, “Poésies...Lais, motets, ballades, rondeaux, et virelais” [dated 1350-1355], f. 189 (persistent URL: <http://gallica.bnf.fr/ark:/12148/btv1b8449043q/f384>).

and differences of presentation then and now promise to remind students of the corollary skills they will need to deal with such materials.

Performers too are already benefitting from direct access of the original versions of music they have hitherto known only through modern editions. Via a link to the Zoomify interface to the DFG³ Viewer, the scans of his autographs at *Bach Digital* offer a particularly hand-some complement to a recent recording of the Bach unaccompanied suites for violoncello (BWV 1007-1012) by my colleague Christopher Costanza of the St. Lawrence String Quartet.⁴ At http://www.bach-digital.de/receive/BachDigitalSource_source_00001200;jsessionid=EBAC8C38-55E50F5F9B2CDD7DF4BB39FF we see the movements of each work as they come from the hand of the composer's wife Anna Magdalena between 1727 and 1731.

IA.2. Digital recovery and enhancement

A systematic examination of the many values of virtual access to original sources would need to focus on more difficult undertakings that involve the manipulation of digital images in order to recapture content that has faded from the page, to enhance it, to facilitate comparison between two prints for "the same" work, to identify copyists, and to retrieve watermark information. Projects have been started in all of these areas. Several have faltered along the way, mainly for lack of continuing funding. The low-hanging fruit of straightforward reproduction has delivered such a bountiful harvest that it seems ill-willed to suggest more intricate processes. The second-level tiers of research considered below all require substantially more funding than simple scanning because the scanning is still necessary, but it only brings the project to the threshold of refining new technologies (mainly associated with graphical analysis, which until *c.* 1990 was largely the province of intelligence networks). Since musicology attracts little support from industry, the adaptation of these technologies is, and is likely to remain, dependent on academic research enterprises.

Digital recovery involves retrieving from a physically corrupted source visual information that is no longer apparent to the naked eye. In music a particularly ambitious and long-sustained project has been the Digital Internet Archive of Medieval Music ([DIAMM](#)), which was organized (by Andrew Wathey, Margaret Bent, and Julia Craig-McFeely) in the late Nineties under the auspices of a series of institutions and sponsors including the Royal Holloway College, University of London; All Souls' College, Oxford; and the British Library. It is now a collaboration of the Music Faculty and the Bodleian Library, Oxford. The fundamental aim of DIAMM has been digitizing all extant musical manuscripts from the later Middle Ages and Renaissance at three resolutions (low for thumbnails and indexing, medium for scholarly consultation, high for questions requiring a closer look at scribal detail). In parallel with the scanning program, an index of relevant collections and their contents has been established.⁵ The website is set up in such a way that users may annotate individual sources. The highest resolution images for the most celebrated sources (e.g. Oxford Bodleian

³ Deutsche Forschungsgemeinschaft, a major funder of German research projects in musicology.

⁴ Bach Digital, like most of the sites mentioned herein, offers a comprehensive site for research. The Bach-Archiv, Leipzig, provides the portal. Access to autographs, prints, and associated libretti for works by J. S. Bach and the extended Bach family are provided together with extensive bibliographical information. The link here is to the manuscript in the Staatsbibliothek zu Preussischer Kulturbesitz, Mus. Ms. Bach P 269.

⁵ The critical apparatus developed for the project is more extensive than the archive of images; that is, many sources currently hold textual material and await scanned images. Searches can be filtered for the presence or absence of images.

Ms. Canonici 213 and Bologna Conservatory MS Q. 15) are also intended to facilitate the production of facsimile prints.⁶

The archive has evolved over time and now includes (for works that have been scanned) musical incipits, full transcriptions of texts within the musical source, and single-folio images with thumbnail views. Metadata fields indicate the genre, number of voices, concordances, and ordinal position within the source. Clefs and mensuration signs are mentioned in the “text” section. An experimental browser facilitates search by location. Two aspects of DIAMM make it particularly appealing: the many points of access represent a near-optimum organization of linked textual and musical (mainly graphical) information and, in selected cases, the high-resolution imaging restores to a useable (virtual) state manuscripts that have been represented in printed catalogues as being lost, irretrievable, or damaged beyond repair. We featured several such examples in a 2001 publication. One is reproduced as Figure 1.⁷

Another lofty goal of DIAMM has been to facilitate digital restoration of material that is faded or otherwise damaged, for experimentation with contrast and color separation by various technical means offers an “inner vision” of content that is invisible to the naked eye. In projects of a similar nature in Italy musical notation has been recovered from what are superficially account books and other unlikely material, in each case concealed in library inventories by the apparent nature of the material. An instance of this in the DIAMM project occurs in the palimpsest Oxford, Corpus Christi MS 144. We see the first and last images (Stages A and D only) of the transit from non-musical “original” to the recovered music.



Figure 1. Oxford, Corpus Christi College, MS 144.

Reproduced by kind permission of the President and Fellows, as originally shown in *The Virtual Score* (Note 6.)

⁶ A facsimile of MS Q. 15 (http://www.omifacsimiles.com/brochures/bologna_q15.html) appeared in 2009, and other titles are in preparation.

⁷ From A. Wathey, M. Bent, and J. Craig-McFeely, *The Art of Virtual Restoration: Creating the Digital Image Archive of Medieval Music*, in: *The Virtual Score* (Computing in Musicology 12), 2001, pp. 227-240.

Another myth that has been laid to rest by technology is the long-held belief that two copies of the “same” print (that is, two instances of publication bearing the same date, the same number of pages, and the same printer’s colophon) must be identical in every way. Through some extensive sleuthing with a Hinman Collator and various digital tools, Philip Brett and Jeremy Smith determined that when images of single pages from the *Psaumes* of William Byrd are compared, they do not in all cases coincide. Byrd appears to have had a change of heart about mensuration signs after a failed publication with Thomas Tallis.⁸ In Figure 2 we see through overlays of “the same” page that some hand modifications were made.



Figure 2. Overlays from two copies of the same partbook for William Byrd’s “My Soule opprest with care and griefe” [Psalm 119] from his book of *Psaumes* (1567). Reproduced by permission from The Virtual Score (Note 5).

A current tendency in long-running projects is to consolidate them. A collaboration between DIAMM and the Medieval Music Database is now in formation. MMDB began as the “Fourteenth-Century Liturgical manuscript inventory,” also known as the *Scribe Medieval Music Database* (<http://www.lib.latrobe.edu.au/MMDB/index.htm>). Designed and developed by John Stinson and John Griffiths at La Trobe University, it contains entries for 3,198 works. In many ways *Scribe* is a model of textual, musical, and liturgical organization. The user interface segregates sanctoral and temporal cycles.⁹ In MMDB neume shapes are shown above unmetred transcriptions, which respect the groupings indicated by the neumes.¹⁰ It provides references to other encoded chants with similar melodic content. A first screen of chants for Ascension (temporal cycle) is reproduced in Figure 3.

⁸ See *Computer Collation of Divergent Early Prints in The Byrd Edition*, in: *The Virtual Score* (as above), pp. 251-260.

⁹ Some of this apparatus is modeled on the Cantus Planus database (http://www.uni-regensburg.de/-Fakultaeten/phil_Fak_I/Musikwissenschaft/cantus/index.htm), a structured data-base with a faceted search. It was developed by a study group of the International Musicological Society, with major contributions by David Hiley, and is seated at the University of Regensburg.

¹⁰ *Scribe* had its own printing system for colored neumes, so that the color content of the original sources could be preserved in the web implementation.

In primo nocturno:

Antiphon 1: CAO 2634 Elevata est magni ficen tia tua super coelos, Deus, alleluia.

Source: Perugia, Biblioteca Comunale 'Augusta', Ms 2797, fol. 57r

Ascensio Domini: I-Pa2797, folio 57r. CAO 2634, first found c. 870, F-Pbn17436

E - le - va - ta est ma - gni fi - cen ti - a tu - a su - per coe - los, De - us,

Modern editions of settings of this text: LR 95; WA 147; LA 244

Related Melodies: 66 have a similarity of more than 80%

Antiphon 1: CAO 2634 Elevata est magni ficen tia tua super coelos, Deus, alleluia.

Source: State Library of Victoria, Ms *096.1 R66A: The Poissy Antiphonal, fol. 142v

Ascensio Domini: I-Pa2797, folio 57r. CAO 2634, first found c. 870, F-Pbn17436

E - le - va - ta est ma - gni fi - cen ti - a tu - a su - per coe - los, De - us,

Modern editions of settings of this text: LR 95; WA 147; LA 244

Related Melodies: 66 have a similarity of more than 80%

Antiphon 2: CAO 2410 Dominus in tem plo san to suo, Dominus in coelo, alleluia.

Source: Perugia, Biblioteca Comunale 'Augusta', Ms 2797, fol. 57r

Ascensio Domini: I-Pa2797, folio 57r. CAO 2410, first found c. 870, F-Pbn17436

Do - mi - nus in tem plo san to suo, Do - mi - nus in coe - lo, al - le - lu -

Modern editions of settings of this text: LR 95; LR 232; GT 540

Related Melodies: 8 have a similarity of more than 80%

Antiphon 3: CAO 1195 A sum mo coelo egressio ejus, et occurus ejus, alleluia.

Source: Perugia, Biblioteca Comunale 'Augusta', Ms 2797, fol. 57v

Ascensio Domini: I-Pa2797, folio 57v. CAO 1195, first found c. 870, F-Pbn17436

A - sum mo coe - lo e - gres - si - o e - jus, et oc - cur - sus e - jus, al - le -

Modern editions of settings of this text: LR 95; WA 147; LA 244

Figure 3. Chants (partial list) for the Feast of Ascension (temporal cycle) in the Medieval Music Database by John Stinson and John Griffiths.

IA.3. Scribal analysis

Any number of handwriting-analysis projects have been initiated over the two decades since desktop scanners became available. As in every other area of digital experimentation, no two projects ever have exactly the same goals. Therefore they never have exactly the same approaches either. While it is thus impossible to compare them directly, the *eNote History* project at Rostock University (c. 2004-2006) remains, despite its discontinuation, well worth our consideration. The University Library has a well-studied collection on manuscripts from the early eighteenth century and adjacent decades. Some of the copyists have been identified by careful documentary research. The role of the digital project was to identify the hands within the collection, to seek a way to categorize them and to examine the transmission processes pertinent to the collection.¹¹

¹¹ The original title of the project was “eNoteHistory — Identifizierung von Schreiberhänden in historischen Notenhandschriften mit Werkzeugen der modernen Informationstechnologie.” The University of Rostock is also the location of the eighteenth-century manuscripts evaluated. Principal participants were Andreas Waczkat (musicology), Andreas Heuer (computer science), and Roland Göcke (graphical data analysis), but several others were also involved. The program operated between roughly 2004 and 2006.

Methods were developed to isolate symbols, to place identified symbols in a lexicon, to construct a similarity index of these graphical types, and then to apply the results to a series of new facsimiles from the collection. It was observed that clusters of scribes might share certain aspects of graphical style. In Figure 4 an overlay technique is again seen, but now for a different purpose.

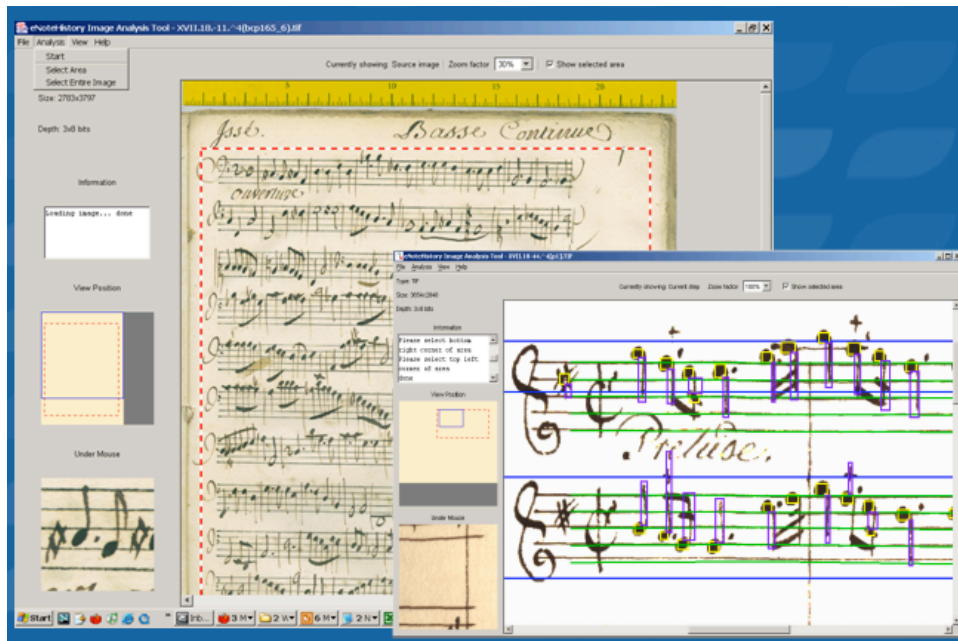


Figure 4. User screen showing source, bounding box, detail of music features, and overlaid images of a short passage in the eNote History project at Rostock University. The music forms part of the eighteenth-century collection of the university library.

IA.4. Watermark Analysis and Lexicon Formation

In the analogue world, watermark studies must rank near the top of the list of undertakings requiring an enormous reserve of patience and persistence. Despite the existence of many reputable catalogues of watermarks, the number of practical difficulties encountered when collating large numbers of similar tracings is very great. The mark may not be visible; the mark and countermark may have become separated; the paper has stretched; the paper has shrunk; the library has no light-box. This litany of problems is overlaid with a human one: no two groups of experts either record the same details about sources or consider the same aspects of watermarks (or handwriting) essential to identification.

Digital photography and image processing have been rapidly pushing aside a century of obstacles and methodological discontinuities. Dexter Edge's project, aimed at Viennese copyists of Mozart, represent an interesting early effort to systematize watermark information. His access to Mozart materials was restricted, but (like many of the digital pioneers) he used the closest available sources—eighteenth-century manuscripts in the library of Louisiana State University, where he was then teaching. In Figure 5, we see a watermark capture in an Adagio and Polonaise for Piano by Iwan Müller.



Figure 5. Dexter Edge, *Mozart's Viennese Copyists*, in: *The Virtual Score* (as above), pp. 261-274. Background image (a "transmissive scan"), the second of four phases of watermark capture, of an Adagio and Polonaise for Piano by Iwan Müller, Louisiana State University Library, Shelfmark 788.6 M88a, as shown in *The Virtual Score*, p. 265. Used by permission.

Within little more than a decade, such efforts have evolved rapidly. Today diverse technical approaches are being pursued in three current German projects pertinent to music. The Wasserzeichen-Informationssystem Deutschland (WZIS) based at Stuttgart's Landesarchiv, is building an apparently graphical search engine to enable users to search by image type. One aspect of the project applies geospatial techniques to overlay European maps with the incidences of specific watermarks. See <http://www.landesarchiv-bw.de/web/50960>.

Wolfgang Eckardt and Steffen Voss made use of the WZIS catalogue during 2008-2010, when they were involved in the digitization of the Schrank II manuscript collection in the Sächsische Landes- und Universitätsbibliothek (SLUB), Dresden (see citations above). A user of the SLUB digital project will find a link to Wasserzeichen which enables him or her to locate all the manuscripts bearing the same watermark. In this way I discovered, *inter alia*, that a Zelenka capriccio (D-DI Mus. Ms 2358-N-4) is one of the two holdings using the same Venetian paper. The steps in the process should be read clockwise from the upper right in Figure 6.

Zelenka capriccio, Venetian paper
D-DI Mus. Ms 2358-N-4.

102	W-DC-102	The lune / "AZ"	nicht bestimmt	nicht bestimmt	Italien	
103	W-DC-103	Wappen Kursachsen zw. Zweigen / "103" 4	Dresden	Schuchart, Johann Gottlob (1717 bis mind. 1739)	Deutschland	ca. 1717 - 1729?
104	W-DC-104	The lune = "A" / Ambrust = "Z"	nicht bestimmt	nicht bestimmt	Italien	
105	W-DC-105	The lune = "ZA" / Ambrust	nicht bestimmt	nicht bestimmt	Italien	
106	W-DC-106	The lune = "A" / Doppeladler	nicht bestimmt	nicht bestimmt	Italien	
107	W-DC-107	The lune / Doppeladler?	nicht bestimmt	nicht bestimmt	Italien	
108	W-DC-108	The lune / "VG"	nicht bestimmt	nicht bestimmt	Italien	
109	W-DC-109	The-lune-Fragment 1	nicht bestimmt	nicht bestimmt	Italien	
110	W-DC-110	Die-lune-Fragment 2	nicht bestimmt	nicht bestimmt	Italien	

Work by Wolfgang Eckardt and Steffen Voss (c. 2008-2010): aimed to catalogue watermarks for all MSS in Schrank II collection

Figure 6. Tracing the watermark for a manuscript in the digitized Schrank II collection of the Sächsische Landes- und Universitätsbibliothek (SLUB).

Upper right: watermark lookup table; lower right image of the watermark.
Upper left: title and first page of first violin part for the Zelenka capriccio in D-DI Mus. Ms 2358-N-4 bearing that watermark.

A third method is being developed in Berlin at the Staatsbibliothek – Preußischer Kulturbesitz (D-B), which reports 52,000 listings in RISM. Gottfried Heinz Kronberger and others are involved.

IB. Textual sources: Natural and structured

IB.1. Full-text and selected text databases

The vast field of text applications contains a great range of resource models. The most common are full-text databases, structured databases, and topic-focused selections of materials. Full-text databases were popular in the 1980s and 1990s for large corpora of core resources. Similar in aim to monumental editions in print, they had the aim to make text (and context) easily searchable. Those interested in exactly what Plato said about music or Aristotle about drama can find the original texts in the *Thesaurus Linguae Graecae*, which has been served in a number of ways at diverse venues. It covers the period from Homer until the fall of Byzantium (1453).¹² Its home site (University of California, Irvine) is at <http://www.tlg.uci.edu/>.

¹² At its initiation (1972) methods for encoding texts in ancient languages were elusive. David Woodley Packard (then a classics scholar, later the founder and director of the Packard Humanities Institute) played a major role in developing systems for both representing and printing in Attic Greek (as well as Hebrew and Coptic for early

A conventional faceted database consisting of textual material is Jennifer Thomas's *Motet Online Database* (<http://www.arts.ufl.edu/motet/default.asp>) at the University of Florida. No full texts are included, but a useful feature is the familiar spreadsheet "rearrange" button enabling materials to be listed alphabetically by text, composer, publisher, et al. Notes give concordances to a wide variety of mostly specialized sources.¹³ At this writing holdings included 18,422 titles from Germany, 11,265 from Italy, 2,201 from France, 1568 from Spain, and 557 from the Netherlands. Thomas is now involved in a collaborative *Chansons et Motets* database,¹⁴ which has a focus on French Renaissance sources but links online editions of individual works from (mainly printed) anthologies of the period. The project is directed by Thomas Schmidt-Beste.¹⁵

An example of a selective database of pertinence to musicology is the *Thesaurus Musicarum Latinarum* (<http://www.chmtl.indiana.edu/tml/start.html>) based on Indiana University. Started and long developed by Thomas Mathiesen and a large contingent of medieval specialists, the TML is now directed by Giuliano Di Bacco. It contains extracts from writings in Latin from the third through the seventeenth century. The TML supports Boolean search for text. It includes encodings of the musical examples found in early writings.¹⁶ It also makes available a very large collection of the illustrations found in them. Given that music was usually analogized in the Middle Ages with celestial phenomena, the graphical corpus is of substantial interest. The TML has spawned similar databases for English, German, and French music theory, which necessarily address a somewhat later chronological period.

IB.2. Digitized libretti

A bonanza of digitized libretti awaits students of Italian opera. The Raccolta Drammatica based at the Ufficio Ricerca Fondo Musicale (online at http://www.braidense.it/cataloghi/-catalogo_rd.php), in which 8,873 texts were listed at this writing, is a model of clarity in linking sources with extensive metadata. Although only four fields are named in the search form, almost any pertinent tag (e.g. the name of a theater, com-poser, librettist, scenographer, singer et al.¹⁷), if placed in the "name" field, may elicit useful responses. The Algarotti Corniani Collection at the Braidense Library, on which this resource is based, consists predominantly of operas but also includes texts of sixteenth-century (prose) *commedie*, of wedding entertainments in the seventeenth, and of oratorios of the seventeenth and eighteenth. It also contains a hearty stock of Rossini, Bellini, and Verdi libretti from the nineteenth century.

A similarly capacious collection is available through an umbrella libretto site (featuring mainly German collections) at the ViFaMusik portal <http://www.vifamusik.de/digital-library/libretti.html?L=1>. Its current total was 7,444 libretti in early June 2013. The largest constituent collection is that of the Bavarian State Library, which includes quite a number of

Biblical texts). A large number of people have been involved in the encoding of the texts themselves and in tools to make them available to a broader public.

¹³ In particular the titles of collections from which the motets come are indicated by the sigla used in Howard Mayer Brown's *Instrumental Music Printed before 1600: A Bibliography* (Cambridge, MA: Harvard University Press, 1965).

¹⁴ <http://lmhs.oicrm.org/chansons-et-motetz/en/database.php>.

¹⁵ Under a grant from the Deutsche Forschungsgemeinschaft, some of its materials will also be available through *DIAMM*.

¹⁶ See the Codes for Noteshapes at <http://www.chmtl.indiana.edu/tml/tofc1.html>.

¹⁷ Especially following Claudio Sartori's *I libretti italiani a stampa dale origini al 1800* (5 vols., Milan, 1990) and its sixteen analytical indices.

German-language translations from French or Italian. An unexpected title of this kind is Richard Wagner's arrangement of Chr. W. Gluck's *Ifigenia in Aulide* (given as *Iphigenia in Aulis* [*Iphigenia in Aulis*], Munich, 1868). The libretto is at <http://daten.digital-sammlungen.de/~db/0005/bsb00054028/images/>.¹⁸

A favorite example (demonstrating the reach of libretti into various aspects of production including in this case choreography) is the libretto for a very early Venetian opera, *Delia* (1639; R 1644). The genre was still so fluid and so closely tied to multifaceted court entertainments that the various elements of the work were assigned to troupes of singers, dancers, and instrumentalists. In the *Azzione* that concludes Act Two, the choreography is provided within the libretto: each dancer (ostensibly a guard in the service of King Admetus) holds a plackard showing a letter or syllable. These being numbered, the libretto demonstrates how the numbers (i.e. dancers) should be shuffled to form a series of three anagrams commending the health of Delia and the power of Venice. See Figure 7.



Figure 7. Consecutive pages describing the Azzione at the end of Act Two of the opera *Delia* (Milan: Biblioteca Braidense, Raccolta Drammatica Corniani Algarotti 453, pp. 57f. Used by permission.)

The text for *Delia*, by Giulio Strozzi, who must have been the choreographer, was set by Francesco Manelli for the premiere performance at the Teatro SS. Giovanni e Paolo in

¹⁸ An arrangement with the original French text (*Iphigénie en Aulide*, Paris, 1774) given in Dresden (1847) and a concert performance on the same text in Zurich (1854) are listed among Wagner's earlier treatments of the same subject, in all cases with major credit assigned to Gluck. The record for a corresponding (partial) print appears in the historical archive of the Bayerische Staatsoper: http://hadbso.bsb-muenchen.de/metaopac/search.do?methodToCall=submit&searchCategories%5B0%5D=-1&searchHistory=&CSId=12933N204S421890dbad01bf-927af3b206e08498d8fd269850&refinTHE=true&refineField=-1&refineHitListName=12_FAST_1952202213-&refineQuery=Wagner+Iphigenie&submitSearch=Suchen&refineType=new. It is useful to note, though, that the material is predominantly in manuscript and that a staging given in Munich in 1816 could not possibly have involved Wagner. [Will Crutchfield reviews a modern revival (1984) of the 1847 Dresden version at <http://www.nytimes.com/1984/07/23/arts/concert-wagner-s-resetting-of-gluck-iphigenie.html>. *Grove Music Online* (http://www.oxfordmusiconline.com/subscriber/article_works/grove/music/O905605pg1#O718699, consulted on 11 June 2013) fails to mention a Munich performance in 1868.]

Venice. The theater was owned by the Grimani family, who were ardent patriots of the Venetian Republic.

IB.3. Festival books

Festival books—miscellanies describing celebrations for a specific courtly or civic occasion—have been largely overlooked until recent times. Great spectacle was a particular stimulus to illustration, and we owe to such sources our knowledge of the grandiose proportions that such occasions could achieve. Digitization suits the heterogeneity of the contents of festival books. Several important collections have been placed online in recent years. Among the largest online collections are those of the Herzog August Bibliothek, Wolfenbüttel; the Staatliche Kunstsammlungen, Dresden; the New York Public Library (Spencer Collection); the Victoria and Albert Museum, London; the Warburg Institute (also London); and the Arsenal and National Libraries, Paris.

One example that conveys so much of the spirit of the occasion is found in a British Library libretto within a libretto: a frontispiece for an (otherwise unknown) serenata called *Proteo con Venere sul Tergo*. The introduction of this lost work appears within a collection of miscellaneous accounts and residues of entertainments¹⁹ for the state visit of Ernst August, Duke of Brunswick-Lüneburg, to Venice in 1685.²⁰ In this British Library reproduction <http://special-1.bl.uk/treasures/festivalbooks/pageview.aspx?strFest=0229&strPage=062> there can be little doubt that Proteus in his omnipotence represents the visiting duke, whose personal pursuit of female pulchritude (Venus) was much noted in annals of that year.

IB.4. Thematic indices

Digital thematic indices are just now coming into view. As with all new approaches, they take a variety of forms. Two fundamentally different repertoires created by Girolamo Frescobaldi (Alexander Silbiger: Frescobaldi Thematic Catalogue Online: <http://frescobaldi.music.duke.edu/>) and Benjamin Britten (Lucy Walker and Jonathan Manton: Britten Thematic Catalogue: <http://www.brittenpears.org/page.php?pageid=590>). The composers lived more than three centuries apart and their repertoires were inevitably very different. In Frescobaldi studies manuscript details are critically important. In Britten studies the diversity of the repertory (ranging from traditional music in practical arrangements to longer works of pungently political cast) is striking. The Britten site is provided with audio examples.

IB.5. Composer legacies

As a model of serving heterogeneous material, the Aaron Copland Collection at the Library of Congress (<http://memory.loc.gov/ammem/collections/copland/>) is difficult to surpass. It contains title and composition lists but is greatly enriched by also making available correspondence, musical sketches, photographs, and draft writings by the composer. It is possible to explore particular tangents in his life (the idea, the sketch, the finished works, the later comments, and in some cases a recording of the same work). A large number of photographs (<http://memory.loc.gov/ammem/collections/copland/phottitindex.html>) give us a sense of the man in all phases of his life, though these are not available for reposting.

¹⁹ *Giuochi festiui, e militari, danze, serenate, machine, boscareccia artificiosa, regatta solenne, et altri sontuosi apprestamenti di allegrezza esposti alla sodisfattione uniuersale dalla generosità dell'A. S. d'Ernesto Augusto duca di Brunsuich ...* (Venice, 1686).

²⁰ The publication, by Andrea Poletti, is from 1686.

The Copland corpus is not necessarily a model for other composer-oriented websites insofar as the texture and length of Copland's life (1900-1990), the variety of his activities, and his infinite zest for new kinds of projects set him quite far apart from many other figures. His lecture on film music (<http://memory.loc.gov/cgi-bin/ampage>, delivered at Columbia University in 1940) is a model of clarity. Milestones in his life are encapsulated in photographs of his public-school class (Brooklyn, 1908); a quiet moment in front of his cabin at the MacDowell Colony (New Hampshire, 1935); a visit by Leonard Bernstein (New York, 1940); a chat with Darius Milhaud (Tanglewood, Massachusetts, 1949); and a banquet at which he was seated near Michael Tilson Thomas (New York or Boston, 1970). These are simply personal selections; myriad paths through the composer's life can be traced with equal ease.

Among numerous other online resources that accommodate composers' documentary legacies in diverse materials, the Franz Schubert site (<http://schubert-online.at>) in the library of the Vienna City Hall deserves mention.²¹ The site offers letters (14 online but this is the tip of an iceberg) and 530 music manuscripts are also present. Watermark information is also available.

IB.6. Music commentary and criticism

Music commentary and criticism from earlier times can be found in a great number of digitized periodicals, though few of them are indexed for music. It is therefore prudent to know something about the periodical and have some provisional dates for the subject of enquiry before proceeding. One of the oldest and most systematically presented is provided by another city library, the Archiginnasio of Bologna. Its Biblioteca Digitale contains most numbers of *Le Gazzette bolognesi* from 1645 to 1796. Entries are organized by place. Two generally unrecognized facts about these assemblies of news (published initially twice a week, later at more frequent intervals) is that entries were organized by place and that the coverage was pan-European, so that among all the similar news aggregations found in the United Kingdom, Spain, France, Germany, Italy and Eastern Europe similar accounts (in locally determined languages) are likely to appear. A reader frustrated by a missing issue in one series may find the information in a parallel aggregation.

IC. Linked reference systems

IC.1. The RISM-OPAC online

The overarching power of digitized materials will be recognized increasingly as more users discover the possibilities of linking several kinds of information pertaining to one source. Recent strides in the online presence of the RISM AII project (manuscript music of the seventeenth and eighteenth centuries) render the preceding half-century of development work a dimly remembered prologue. An illustration will serve the purpose of description (Figures 8a, 8b).

²¹ Collaborators include the Institute of Musicology and Psychology in Cologne and the Institute of Musicology in the University of Vienna.



Figure 8a. Autograph cadenza from Vivaldi's violin concerto in A major (RV 340) as listed in the RISM-OPAC (http://opac.rism.info/index.php?id=6&no_cache=1&L=0&tx_bsbsearch_pi1%5Bquery%5D%5B0%5D=RV%20340&tx_bsbsearch_pi1%5Bid%5D=212000183) from D-Dlb, Mus MS 2389-O-43.

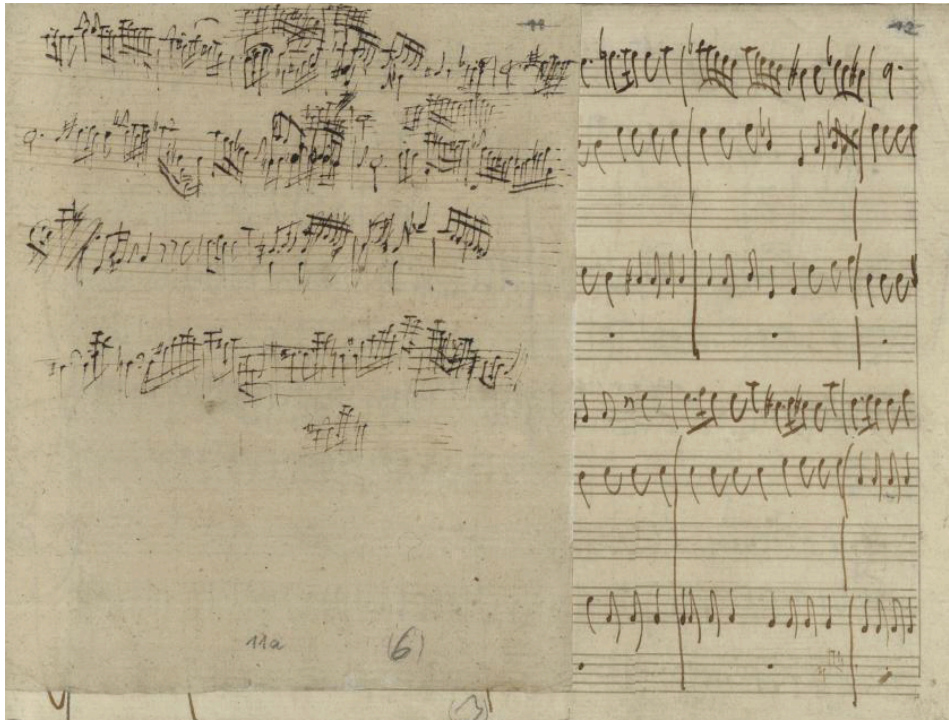


Figure 8b. An alternative elaboration of the second-movement theme, inserted leaf (preceding f. 15), in a digitized copy of the manuscript posted at <http://digital.slub-dresden.de/ppn316139777>. The cadenza for the third movement is appended to the manuscript (f. 28).

This autograph (c. 1716-17) is found in the Schrank II collection in the Sächsische Landesbibliothek, Dresden. The entire collection (3,113 manuscripts of the eighteenth century) is not only available online (with a downloadable Excel spreadsheet describing its contents) but is digitally cross-referenced to the appropriate RISM entries. The home site for the manuscripts is <http://digital.slub-dresden.de/listenansicht/>, but finding it from within the library's many other digital depositories can be trying. The RISM hyperlinks simplify the problem. The pertinent listing here is: http://opac.rism.info/index.php?id=6&no_cache=1&L=0&tx_bsbsearch_pi1%5Bquery%5D%5B0%5D=RV%20340&tx_bsbsearch_pi1%5Bid%5D=212000183. The consistency between RISM and the Hofkapelle is very reassuring: even the cadenza that is scribbled on an inserted leaf is indexed in RISM.²²

²² Under the "Recherche" heading at www.schrank-zwei.de/suche/ one can also find the very valuable *Schreiber- und Wasserzeichen-Kataloge*. For this concerto, which was dedicated to Pisendel (probably during his visit to Venice in 1716-17), the transcriber is of course Vivaldi. The watermarks found on the paper are W-DI 131 and W-DI 394, neither of which (we learn from the watermark database) finds many concordances in the rest of the Schrank II collection.

IC.2. Opera around 1800

A similar hyperlinking effort, in this case wedding libretti and metadata to digitized opera scores, is *Die Oper in Italien und Deutschland 1770-1830*, based at the universities of Cologne and Bonn (<http://www.oper-um-1800.uni-koeln.de/>). It links in 125 opera manuscripts from the Sächsische Landesbibliothek in Dresden to several other kinds of information in a structured database (http://www.oper-um-1800.uni-koeln.de/alle_gescannten_opern.php?anfangsb=Dresden&ordnen=sammlung).

II. Musical scores and their analysis

Musical scores are neither fish nor fowl—neither purely graphical nor inherently sounding. To paraphrase a wonderful remark by Margaret Bent, “Music exists only in sound, but paradoxically, sound is its least stable element.”²³ The instability of performance time complicates all efforts to synchronize scores with sound. We turn to scores for a stable record of what a composer wrote and of what we should perform. As with the TLG, it is impossible to search material that has not been encoded. A major effort of projects initiated in the 1980s was to create full-text “databases” consisting of complete works by major composers, so that they would be searchable in the same way as texts were then searchable.

IIA. Scores and parts scanned from early printed sources

IIA.1. The Danish National Library Sheet Music Archive

Collections of early printed music available for consultation are under rapid development. The Danish National Digital Sheet Music Archive is possibly the most systematic. Found at <http://www.kb.dk/en/nb/samling/ma/digmus/index.html>, it contains some manuscripts and libretti but principally serves scores. A monthly bulletin alerts users to new digital accessions. It also has an archive of theatrical source material. Although holdings of Danish materials are highlighted (all works of the Carl Nielsen edition can be downloaded without cost), digitized materials include large quantities of other European sources.

IIA.2. Early Music Online

Another robust site is Royal Holloway College’s Early Music Online (<http://www.early-musiconline.org>), which contains 327 printed anthologies of music from the sixteenth and seventeenth centuries. This collection consists of part-books rather than modern editions. Italian prints are especially well represented. It will be complemented by 1800 music prints of the sixteenth and seventeenth centuries under development by ViFaMusik. Here too anthologies (a popular kind of publication in the sixteenth century) play a significant role and part-books will constitute much of the material made available.²⁴

²³ M. Bent, *Editing Early Music: The Dilemma of Translation*, in: *Early Music* 22/3 (1994), p.391.

²⁴ See <http://vifamusik.wordpress.com/2012/01/26/die-notendrucke-des-16-und-17-jahrhunderts-mit-mehrstimmiger-musik-in-der-bsb-digitalisierung-und-online-bereitstellung/>.

IIA.3. Sibley Digital Scores

A rich source for nineteenth-century performing materials can be found, under the auspices of the Eastman School of Music, under the title Sibley Digital Scores. The digital holdings now exceed 21,000 public-domain scores. Most are printed but some manuscripts are also included. The site is maintained by the University of Rochester (Eastman's host institution) and can be accessed here: <https://urresearch.rochester.edu/viewInstitutionalCollection.action?collectionId=63>.

More specialized collections are also numerous. University of North Texas has digitized a collection of scores and manuscripts of the music of J.-B. Lully (see <https://urresearch.rochester.edu/viewInstitutionalCollection.action?collectionId=63>).

American sheet music of the nineteenth century is represented by numerous collections, mainly in the Western U.S.²⁵ The University of North Carolina has a particularly powerful search engine (<http://dc.lib.unc.edu/cdm/search/collection/sheetmusic>) for its collection.

IIB. Repositories of encoded Musical materials

Online repositories of digital music sources are useful in countless ways: they enable us to see primary sources that are otherwise unavailable; to examine handwriting and scribal features; to identify watermarks; and to compare one source with another. In these cases “searchability” means retrieving descriptive information from metadata. Digitized images of graphical material are not in and of themselves searchable.

IIB.1. MuseData

Encoded scores can be used to facilitate new editions, comparative editions, the production of performing materials to almost any specification, and the excerpting of material. They facilitate myriad kinds of analysis, some conventional, some newly devised to take advantage of the availability of digital data. These ideas lay behind the establishment of the *MuseData* project (<http://www.musedata.org/>) of the Center for Computer Assisted Research in the Humanities at Stanford University. *MuseData* has myriad component parts—the original data (encoded either from out-of-copyright editions or from newly commissioned ones); translations of the data for diverse purposes; scores (sometimes with performing parts) in PDF format; and one particular translation (to the Humdrum Toolkit kern format) intended for analytical tasks.²⁶

Because users seem to find data *per se* intimidating, the Center has been creating a series of websites for widely recognized works among its holdings. Examples include Beethoven's symphonies (<http://www.ccarh.org/publications/beethoven-symphonies/>) and string quartets (http://wiki.ccarh.org/wiki/Beethoven_String_Quartets), G. F. Handel's *Messiah* (<http://wiki.ccarh.org/wiki/Messiah>), and the published *oeuvre* of Corelli (http://wiki.ccarh.org/wiki/MuseData:_Arcangelo_Corelli). All include scores and parts but also provide some historical context and information on performance issues.

All projects relating to encoded musical data remain in a pioneering, quasi-experimental state, because there are no widely recognized best practices for encoding data. Music lacks an ASCII code, partly because it is enormously complex but partly because every encoding

²⁵ See <http://www.lib.unc.edu/dc/sheetmusic/links.html>.

²⁶ The encoding system was designed and all the software for it developed by Walter B. Hewlett. Craig Stuart Sapp has written a large proportion of the translation software and analytical routines.

system favors one kind of music or another. In combination with the high degree of heterogeneity in repertoires across time and space, the general view of specialists in the area is that a one-size-fits-all encoding scheme is impractical. It will make some tasks easier but others more difficult, and the data will be so dense that program debugging will be all but impossible. In the absence of a normative approach, the largest collections of encoded materials are pegged to Common Western Notation, i.e., a reliance on conventional barring, regular rhythmic values; pitch encoding that includes single and double sharps and flats; support for text underlay in languages based on the Roman alphabet; and a reasonable palette of tempo indications, articulation marks, ornaments, and so forth—the component parts of scores written from the seventeenth through the mid-twentieth centuries.

IIB.2. Humdrum kern

Data is the CMN encoding format used with the *Humdrum Toolkit*,²⁷ which was developed in the 1980s by David Huron. Several other researchers have expanded the tools in subsequent years. Here at Stanford we offer a course on musical analysis using Humdrum. It is based on encoded materials at the KernScores website: <http://kern.ccarh.org>. A syllabus and some instructional materials can be found at http://wiki.ccarh.org/wiki/Music_254/CS_275b_-_Spring_2013_Syllabus.

Those interested in exploring *Humdrum* may wish to visit CCARH's Humdrum Portal (<http://humdrum.ccarh.org/>) and consider joining the **hug [*Humdrum* users] mailing list. The original Humdrum site (<http://humdrum.org/HumdrumDownload/downloading.html>), an extensive reference manual, and pedagogical materials (<http://humdrum.org/Music824-/music824.html>) can be found at the Ohio State University.

IIC. Search tools for encoded data

Search programs are often based on a data format optimized for search efficiency. Although there is no common format or instantly useable “neutral” format for encoded data,²⁸ direct data translation has often proved to be preferable. *Humdrum* supports neutrality: it can withstand the stripping of data unnecessary for a specific query.²⁹

IIC.1. Themefinder

Humdrum's uses are so diverse that it is impossible to corral them in a few short sentences. Our longest standing tool, *Themefinder* (1996), continues to be put to frequent use interactively, especially to determine what search strategies are best suited to particular situations. The data samples it contains (music from the seventeenth through early twentieth centuries, selected Renaissance music [mainly motets], and folksongs [mainly German]) are

²⁷ This requires particular emphasis, because *Humdrum* is designed to work with “just enough” data—simple transcriptions of pitch only, rhythm only, and non-Western music of myriad kinds.

²⁸ The most common defaults to which users resort are MIDI files (for sound and for import to notational software) and MusicXML. MIDI and MIDI-based tools are vulnerable to inaccuracies in rhythm and enharmonic pitch rendering. MusicXML, which was originally a synthesis of formats used in MuseData and *Humdrum* encodings, is now (from 2011 onward) a proprietary format subject to alteration by the company succeeding to the management of *Finale*[®]. Until recently the public company was MakeMusic, which has now been taken private for a consortium of investors.

²⁹ Music21, developed by Michael Cuthbert and others, is a more recent set of tools for music query. It has many virtues, though it is less well documented than *Humdrum*.

far less diverse than we might wish. Yet by combining these public components with some privately held ones we were able to use a base of 100,000 music incipits to investigate the relative efficiency of different strategies for music search (2004).³⁰

Themefinder (<http://www.themefinder.org>) offers a less-more continuum of pitch detail in the query description as well as filters for key, meter, and mode. See Figure 9.

The image shows the Themefinder search interface. It features several input fields and dropdown menus for search criteria. On the left side, the fields are: Repertory (set to 'Classical'), Pitch, Interval, Scale Degree, Gross Contour, Refined Contour, Location (with radio buttons for 'beginning of theme only, or' and 'anywhere in theme'), Key (set to 'Any'), Mode (set to 'Any'), and Meter. On the right side, there are help icons and explanatory text for each field: 'type of music to search', 'A-G, sharp=#, flat=- e.g. C E- G F#', 'maj=M, min=m, aug=A, dim=d per=P, fifth=5, up=+, down=- e.g. +m9 -P8 +M3 P1', 'do=1, re=2, mi=3, fa=4, so=5, la=6, ti=7 (mode insensitive). e.g. 34554321', 'up=/, down=\, unison=- e.g. //\-/ or uudsu', 'up step=u, up leap=U, down step=d, down leap=D, same=s. e.g. uUDsdu', and a '?' icon. At the bottom left, there is a 'Submit Search' button highlighted with a red box.

Figure 9. The Themefinder search interface (<http://www.themefinder.org>).

Anecdotal evidence suggests that most users do not remember a piece of music with sufficient precision to prefer a pitch-based search. Intervallic search is more forgiving of error but the queries are somewhat cumbersome to construct. The scale-degree search is the simplest and easiest to construct. All of these search-types admit wildcard characters for fuzzy searches. Users may also select whether the query should match the starting notes of an entry or not. The mode filter is more useful when the material sought is in a minor key, because more than 80% of the holdings are in a major key. Counter-intuitively, music of uncertain attribution is best pursued through a gross-contour search, because it seems that anonymous pieces are somewhat more prone to mis-transcription than those of known authorship. What our 2004 study showed was also counter-intuitive: the shortest route to a discrete match, in a sample of 100,000, was achieved by using the combinatory measure gross pitch contour (as above) *plus* gross rhythmic contour (fast-slow-same level of discrimination of durational values). It surpassed a series of progressively more refined pitch-only searches.

³⁰ See C. S. Sapp, Y.-W. Liu, and E. Selfridge-Field, *Search-effectiveness measures for symbolic music queries in very large databases*, in: *5th ISMIR, Barcelona, Spain in October, 2004*. Downloadable from <http://ismir-2004.ismir.net/proceedings/p051-page-266-paper135.pdf>.

IIC.2. Other music-search tools

Apart from music21,³¹ other search-enabled projects may concentrate on text search, or an intermediate layer of information that is invisible to the user. The multi-faceted *Liber usualis* project (<http://ddmal.music.mcgill.ca/>) at McGill University does not fit neatly into this category, since it provides an array of materials and strategies. At this juncture we are too inexperienced to know whether searching by neume shape is more or less useful than searching by chant text or liturgical function, although the Medieval Music Database makes this question easy to investigate.

Peachnote is a worthy contender in the general space of quantity of material. It has the strength that a best reading is culled automatically from multiply-scanned prints combined with the weaknesses that the accuracy of encoded material is not humanly verified and the search algorithm, being MIDI-based, is inadequate for enharmonic pitch discrimination (as one will notice in the thematic quotation from Beethoven's Ninth Symphony on its home page: <http://www.peachnote.com/#!nt=singleNoteAffine&npq=62+0+1+2+0+-2+-1+-2>). The Google timeline showing thematic concordances from before 1700 is meaningless, since the conditions of tonality and temperament rested on assumptions that are different from those that underlie MIDI files.³²

IIC.3. Visual score search: The Josquin Research Project

The Josquin Research Project, directed by Jesse Rodin, has a number of purposes. Among these the overriding aim is to evaluate the legitimacy of authorship claims. Since the project is still in progress, it is not possible to evaluate it or to compare its results to analogue efforts for answering similar questions, but only to show a few of its novelties. JRP (<http://jrp.ccarh.org>), which is being implemented by Craig Sapp, currently contains roughly 600 pieces and/or mass sections of sacred and secular music from roughly the time of Josquin. Its search and analysis tools are self-explanatory (Figure 10). Its files can be downloaded.

³¹ <http://web.mit.edu/music21/>.

³² The defect of MIDI files as a basis for fine-grained searching of notated musical content is that they do not distinguish between C# and Db (or between other name pairs for accidentals). Over time some algorithms have proved better at discriminating through contextual reasoning, but in general MIDI-based search accuracy is much higher for simple music in major keys than for repertories that are complex, in minor keys, pre-tonal, or post-tonal.

The Josquin Research Project

Project Director: Jesse Rodin, Stanford Music
Technical Director: Craig Sapp, CCARH

WORK LIST

Go to: [Agricola](#) | [Busnoys](#) | [Du Fay](#) | [Josquin](#) | [La Rue](#) | [Mouton](#) | [Obrecht](#) | [Ockeghem](#) | [de Orto](#)

No. Title (click title for more info.) S Scores MIDI Voices

Josquin des Prez (289 of 336 works)

Masses (31)

No.	Title	S	MIDI	Voices
1201	Missa Ad fugam	S	MIDI	4
0701	Missa Allez regretz I	S	MIDI	4
0301	Missa Ave maris stella	S	MIDI	4
0304	Missa Cum iocunditate	S	MIDI	4(-5)
0302	Missa Da pacem	S	MIDI	4
0303	Missa De beata virgine	S	MIDI	4-5
0903	Missa Di dadi	S	MIDI	4
0703	Missa D'ung aultre amer	S	MIDI	4
0801	Missa Faisant regretz	S	MIDI	4
0401	Missa Ferialis	S	MIDI	4
0802	Missa Fortuna desperata	S	MIDI	4
0402	Missa Gaudeamus	S	MIDI	4
1101	Missa Hercules dux Ferrarie	S	MIDI	4(-6)
0601	Missa L'homme armé quarti toni	S	MIDI	4
0602	Missa L'homme armé sexti toni	S	MIDI	4(-6)
0603	Missa L'homme armé super voces musicales	S	MIDI	4
1102	Missa La sol fa re mi	S	MIDI	4
0501	Missa L'ami Baudichon	S	MIDI	4
0901	Missa Malheur me bat	S	MIDI	4(-6)
1001	Missa Mater patris	S	MIDI	4(-5)
9901	Missa Mi mi	S	MIDI	4
1002	Missa Missus est Gabriel Angelus	S	MIDI	4

SEARCH TOOLS

Pitch

Interval

Rhythm

Sort by:

Accidentals:

Composer:

Genre:

Mensuration:

[instructions](#)

ANALYSIS TOOLS

Repertory:

Thematic index

Vocal ranges

Parallel octaves/fifths

Measure-long rhythmic patterns

Attacks per measure

Figure 10. Search and analysis page for the Josquin Research Project at Stanford University.

IID. Analysis tools for encoded data

IID.1. Visualizing musical texture and form: Malinowski's Music Animation Machine

An intermediate space on the continuum between public use and academic value is occupied by more broadly focused visualization tools, such as Stephen Malinowski's *Music Animation Machine*. Malinowski depicts formal architectures in the process of their own evolution. His work, from 1986 through the present, is a dynamic art form in its own right, despite its obvious dependency on the underlying work of others. In their accumulation, Malinowski's videos show the seemingly infinite diversity of textures and structural devices used by composers from the sixteenth through the twentieth centuries. Video platforms such as Vimeo and YouTube have greatly increased public appreciation of his work. An early mural gives thumbnail impressions of keyboard works by numerous composers (Figure 11).

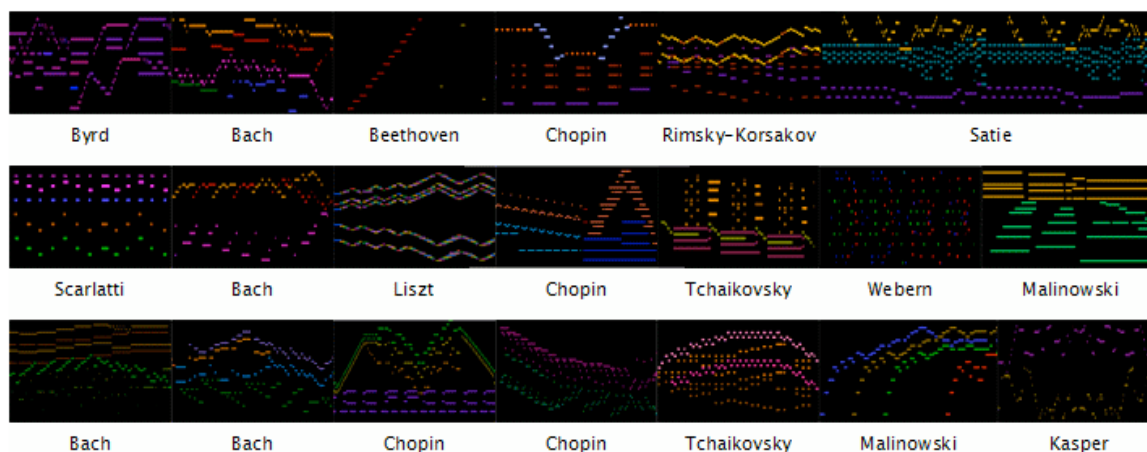


Figure 11. A series of spliced segments of real-time “piano rolls” of individual keyboard works in Stephen Malinowski’s Music Animation Machine.

Malinowski’s video of the Hallelujah Chorus of Händel’s *Messiah* (http://www.youtube.com/watch?annotation_id=annotation_749068&feature=iv&src_vid=Htkx5LNm3SE&v=JtoNHnR_WhE, 2009) brought particular recognition to his work. It is cast in the traditional MAM mode, with voice parts differentiated by color, duration by width, and a white cursor overlay to keep the viewer in synch with the recording. His rendering of Stravinsky’s “Rite of Spring, Part I” (<http://www.youtube.com/watch?v=02tkp6eeh40>, 2012) is MAM’s latest display. It and its analogues were released for the work’s centennial (2013). In this case diverse shapes identify timbres, colors pitches, and sizes dynamic level. The complexity of textures is particularly easily perceived in this neon-colored video.

IID.1. Visualizing harmonic structure: Sapp’s Keyscapes

Beyond search in the simplest terms, the greater power of encoded data is in facilitating accurate and efficient analysis across whole corpora of encoded music. However, the reporting of results in ways that are suitable for those not trained in statistics still awaits consideration. In this respect the analysis of harmonic structure in movements and works by way of the visualization routines labeled “keyscape” in the doctoral thesis of Craig Sapp (2011) provide the basis of endless comparison.³³ A keyscape (Figure 12) shows elapsed time (horizontal dimension) and level of generalization of functional harmonic usage (vertical dimension). The temporal resolution is most precise along the bottom of the figure, most general at the top.

³³ C. S. Sapp, *Computational methods for the analysis of musical structure*, Ph.D. dissertation, Stanford University: <http://purl.stanford.edu/br237mp4161>.

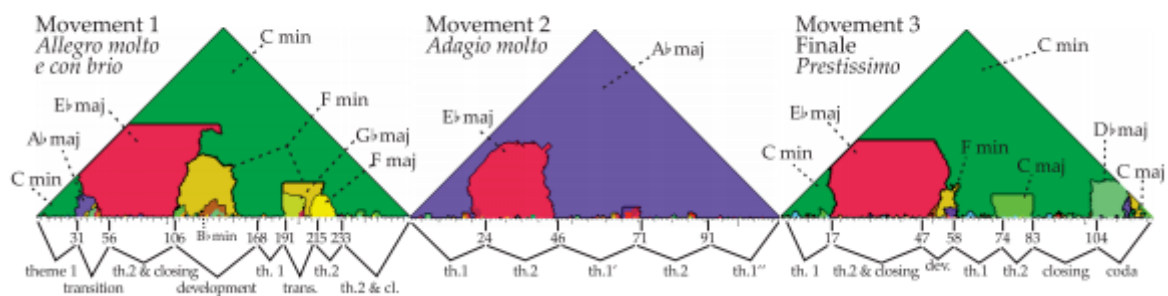


Figure 12. Harmonic structure of Beethoven's piano sonata Op. 10, No. 1 in C Minor.
 Color code: Green=C minor; Scarlet=Eb major; Purple=Ab major.

Users unfamiliar with such techniques will no doubt require time to become accustomed to synoptic views such as this one. Yet the value of such summarization techniques in studying changes over time (or place) in large repertoires of encoded pieces is beyond dispute. As one example, Beethoven's harmonic inflections in each movement of the piano sonata Op. 10, No. 1 in C Minor can be compared above. (The middle movement is in Ab Major.)

III. Sounding sources

The needs of musical research surpass those of all other disciplines in their requirement that audio, text, and graphics ultimately be linkable. The importance of sound to musicological study is obvious. The technical requirements of audio vis-à-vis text and graphics formats are not at all obvious. Sound happens in time; space necessarily does not. While this fundamental incompatibility of sound and graphics has been superseded by various hidden intermediate schemes bilaterally linked to each, robust systems suited to any user will not be easy to provide.

Audio (and video) formats evolve at a rapid pace, perplexing library budget directors with an unending need to migrate materials to newer media, in order that their content not be lost. A helpful redress is delivered by the internet, where those with music on obsolescent media have repeatedly found ways to recreate listening experiences through various schemes to render the music audible.

IIIA. Access facilitated by the internet

The distribution of sound materials of many kinds is greatly constrained by copyright questions. Those repertoires cited below become viable either because their holdings were never subject to copyright or because the holders own all the rights to the material. For many of the projects mentioned above sound output is not a legally viable option. Among the best known sound archives with an online presence are the Library of Congress, the Smithsonian Institution, the British Library, several university sound archives (Stanford and Yale prominent among them), and various European counterparts. At least six sites serve Thomas Edison Diamond Disk reproductions. Some archives have one predominant repertory (e.g.,

Rogers and Hammerstein at the New York Public Library). These are merely representative of the breed. Many others continue to seek ways to provide more audio materials online.

IIIA.1. Internet radio and library sound archives

Internet radio also offers countless opportunities for listening to music from all over the world. Special collections, such as Alan Lomax's 17,000 recordings of folk music, are increasingly finding their way onto the web. (The Lomax collection is included here: <http://research.culturalequity.org/home-audio.jsp>.) Internet radio stations devoted to specialized repertoires (Indian sitar, sarod, and tabla music, for example) reside in a vast virtual space. Broadcasting services such as the BBC amassed huge repositories of both recordings and unpublished editions, but little of this is currently available in digitized form.

Among library sites, the British Library's "Sounds" includes some substantial items (such as major symphonies of the nineteenth-century) that can be heard online. Please visit <http://sounds.bl.uk/Classical-music>.

At our institution, the Riverwalk Jazz collection (4000 song performances: <http://riverwalkjazz.stanford.edu/>) has recently been made available online via internet radio. The copyright statement reads, "Due to copyright restrictions, we are unable to provide a jukebox-style program stream where the user selects individual radio shows. It's like a box of chocolates. Drop in and discover what's on the air." Cataloguers have had to be ingenious to facilitate structured access to streaming material (juke-box [random] access violates US copyright).

IIIA.2. Folk music archives

Its "open culture" home operates on a different principle from the current instantiation of the Deutsche Volkslied Archive (DVA; see <http://www.dva.uni-freiburg.de/>), which offers extensive look-up facilities for titles, lyrics, and editions of German folk music. The Archiv, based at the University of Freiburg, was established in the 1920s. It has taken full advantage of opportunities for dissemination on the internet. Similar in general aim but significantly different in implementation is the Austrian Volksmusikdatenbank, which has even earlier origins (1904) but is less centralized and links to folksong archives in individual states. (See <http://www.volksmusikdatenbank.at/>.) It also incorporates programs of events and promotes opportunities to learn about folk music. Whereas the role of sound in the Lomax collection is paramount, it is more muted in these European collections.

What is noteworthy about the two European projects is that non-digitized portions of both archives include large repositories of typescript code (for music transcription) that anticipate what is known today in digital musicology as the Essen Associative Code (EsAC). Converted to this transparent encoding format, the music becomes eminently searchable. In fact numerous programs were written in the period 1982-1996 by and under the direction of Helmut Schaffrath, who was affiliated with the Essen Hochschule für Musik.³⁴ (No active program for digital rescue of typescript "data" seems to exist today. This is regrettable because well over 100,000 folk songs appear to have been transcribed in this uniform manner in various locales of Central and Eastern Europe.)

³⁴ Dr. Schaffrath died prematurely in 1994. His students attempted to carry on his work for a few years, but without institutional support, this proved to be difficult. A prototypical version of the EsAC code (and some of the Essen programs) in the earlier mainframe-based work of Dr. Wolfram Steinbeck (now a leader of the *Oper um 1700* cited above) at Kiel University. See the Habilitationsschrift *Struktur und Ähnlichkeit: Methoden automatisierter Melodienanalyse* (Kassel, 1982).

IIIA.3. Special media

The New Zealand Pianola site

(http://www.pianola.co.nz/rollscans/rollscan_search_results.asp) is a model of transparency and efficient access. It is the product of a long development cycle that started twenty or more years ago with a determination to mine piano rolls and convert the roll data to a playable digital format. Numerous technical hurdles had to be overcome. The Rollscan site is rich in repertoires popular between 1885 and 1935 (the age of “parlor songs,” the heyday of ragtime, and the peak years of player piano use). Some art music is also found, including Edvard Grieg’s musical impression of a “Butterfly” (Op. 43, No. 1, 1886), captured from a Welte-Mignon “reproducing record”: [http://www.pianola.co.nz/rollscans/midi/Butterfly\(1886\)_-Welte-MignonReproducingRecord-C1275_TheComposer\(EdwardGrieg\).mid](http://www.pianola.co.nz/rollscans/midi/Butterfly(1886)_-Welte-MignonReproducingRecord-C1275_TheComposer(EdwardGrieg).mid). The Americana found at this site is so abundant that it is often possible to compare two or three renditions of the same piece.

The nuances of individual performers are easily evidenced on account of the high quality of modern reproductions. This 1931 recording of “As Time Goes By” ([http://www.pianola.co.nz/rollscans/midi/AsTimeGoesBy\(1931\)_Ampico-213801_VictorArden.mid](http://www.pianola.co.nz/rollscans/midi/AsTimeGoesBy(1931)_Ampico-213801_VictorArden.mid)) comes from an Ampico recording. It features Victor Arden, whose style was characterized as “pyro-technical.”

IIIA.4. YouTube

In certain corners of global repertory, YouTube embeds in its video service countless treasures—that is materials that cannot be found in libraries or archives because the repertory is in some respect too remote. The site is, of course, uncurated. There is no meaningful metadata, and one has to take at face value the information provided by the uploader. Despite the availability of some large commercial collections of materials relevant to ethnomusicology, YouTube sometimes offers more. Finding it may be difficult, but a compensation of the associated video is that it may convey better than any book or record the ordinary context in which a certain kind of music typically occurs.

Since YouTube is not in any musicology syllabus, I cite an area of value from recent experience. In 2011 I needed to come up quickly with content relevant to music in a Third World society. I chose Samoa because of personal familiarity. When I visited the South Pacific decades ago I searched in vain for recordings or transcriptions of indigenous music, but none was to be found. In subsequent years I met a few ethnomusicologists (based in New Zealand and Australia) who did detailed research on particular repertoires. They published valuable studies, but of course the studies are silent. One of the things that fascinates Westerners about Samoan music is the universal aptitude Samoans (and Tongans) have for spontaneously harmonizing in four voices. This skill is undoubtedly tinged by a missionary past, but the utter dependence on a spontaneous oral tradition leaves little room for any kind of formal system lying beneath the surface today. Samoans value their cultural past, and YouTube abounds with materials suitable for purposes such as mine. This performance of the Samoan chant “Savalivali maia” (“Go for a walk with me”: http://www.youtube.com/watch?v=bKA8NW_Tuvw) represents a broad horizon of comparable possibilities.

IIIB. Sound Analysis

IIIB.1. Performance snapshots

Apart from general support for the exploration of all kinds of music, the combination of tools to represent various parameters of musical sound by visual means has given rise to the fine art of performance analysis. The standard approach of many projects is to compare multiple recordings of the same work to study individual differences in interpretation and expression. Diverse graphical means have been devised to report the comparative results. In Gerhard Widmer's lab (Austrian Research Institute for Artificial Intelligence, Vienna; with major contributions by Simon Dixon and Werner Göbel) *performance worms* (Figure 13a) show differences of tempo and dynamics in an overall environment of temporal decay. See <http://138.37.35.209/people/simond/pub/2002/icmc.pdf>. Many subsequent projects in the same lab have explored relationships between cognition and listening.

IIIB.2. The Chopin Mazurka Project

In CHARM's Chopin Mazurka Project in the UK (Royal Holloway College, University of London, 2005-2007), led by Nicholas Cook and John Rink) Craig Sapp's *timescapes* (Figure 13b) offer a synoptic view of tempo variation. See http://www.charm.rhul.ac.uk/analysing/-p9_4.html. By analogy with keyscares (above) they compared elapsed time in a given performance with a reference score to examine differences between two performances.

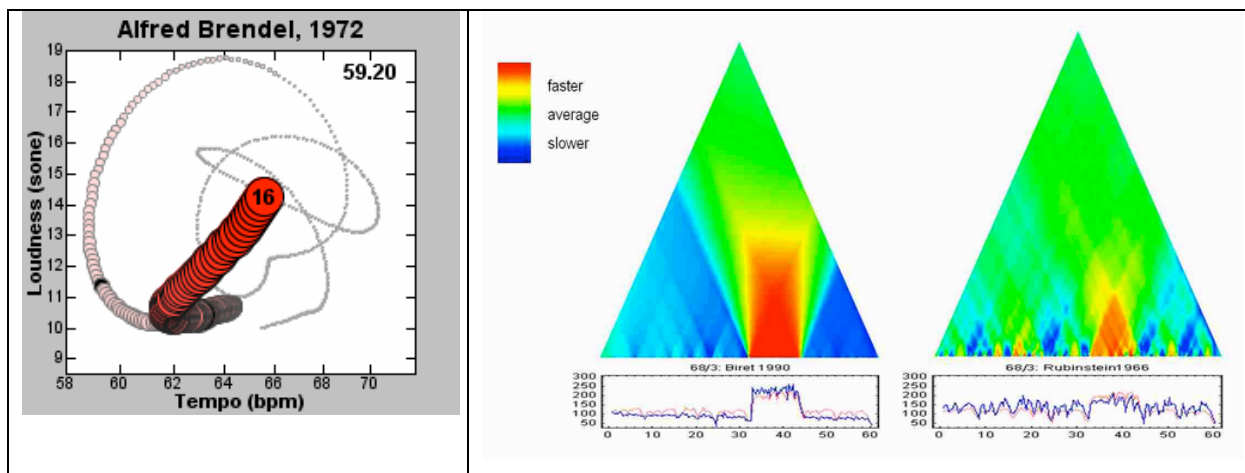


Figure 13a. Performance “worm” giving a correlated view of tempo and loudness in relation to elapsed time.

Figure 13b. “Timescapes” comparing two performances (Biret, 1990, and Rubinstein 1966) of Chopin’s Mazurka Op. 66, No. 3.

IV. Reflections on authority

Having surveyed representative examples of the wealth of graphical, textual, and audio evidence that has arrived at our fingertips, where do we go from here? Partly by its rising quantity, we are forced to focus on three main areas of need in the future: (1) management of digital resources, (2) improved access to non-Western materials, and (3) further exploration of music search.

The need for adequate management seems obvious, but maintaining meta-information about digital resources is far from easy. Resources grow, but their locations also change.

Projects lose funding and “go dark.” Users fail to find what they are seeking, or even what they found a year or two ago. It is easy to lose heart with evaporating sources. We need a permanent repository of “last known site” links. Permanent archives for all projects that achieve a web presence only to lose it again are badly needed, as are archives for curated collections of data in any digital form. This need has been perennially discussed across all academic disciplines for more than twenty years, but no archiving scheme seems to survive for more than three years.

The case of non-Western materials is also self-evident, but solutions are harder to achieve. Many kinds of music for which there is no notation either on paper or online exist only in sound, but among sound files there is no authority and often enough little identifying information. To see a repertory shared by large communities of people transformed from sound to sonograms is both exciting and dismaying: there is some kind of visible record of the music, but it communicates more to engineers than to musicians. Is YouTube destined to be the only bridge between traditional culture and modern science?

In the case of music search, scholars have become accustomed to accepting many substitutes for the real thing. Symbolic search and audio search depend on entirely different methods. Symbolic search can be very specific, but audio search depends on fairly crude determinants, such as beat. Services that project an impression of synchronized search do it by subterfuge in the form of invisible interior mappings between the two. Library metadata fills the gaps for most of the recorded material used in Academe, but this may change as such holdings become more diluted, in the minds of students at least, by offerings scattered across the internet. Symbolic search, on the other hand, is well studied and sophisticated comparisons of diverse approaches have been available for several years. Its utility is somewhat crippled by limited quantities of encoded data. This lapse is primarily a reflection of the disinterest of funding agencies in supporting data development.

Bubbling just beneath the surface of many discussions of digital resources is a nagging question: How do we control scholarly authority in the age of YouTube and Wikipedia? Travesties of long-established knowledge are easy to find. Students are gullible, and the battle against ill-gotten pseudo-information looms large in today’s classroom. Is there any hope of saving future generations from an avalanche of half-truths?

My long immersion in digital precincts has persuaded me that the transparency of the internet may go a long way in redressing this problem. The time it takes to gather knowledge on a particular subject is constantly decreasing. If one contrasts the efficiency of my starting example (a research trip to Bibliothèque Nationale de France, in my case starting from California), there is simply no comparison with the convenience of a few clicks to make a virtual visit to the relevant source material. There are also no distractions: I have a question, I seek an answer (*sans* passport, hotel reservation, or reader’s card interrogation). Most of us can work more accurately and efficiently in our native habitats. But should convenience be our only yardstick?

My belief is that the “authority” question has a reverse side that still lingers on the horizon: Will our students, working from *prima facie* evidence, see (perhaps implicitly) through the received truths that we have been teaching them? We have an ancient scholarly tradition in which each generation is encouraged to “stand on the shoulders” of its predecessors. The giants on whose shoulders we have been standing were themselves dependent in large measure on secondary sources and second-hand information. Only for the past century have some musicologists been regularly able to see selected primary materials. Our students can

view manuscript after manuscript, in scattered physical locations, in a single afternoon. What will this do to scholarship?

It may be that the eternal verities will be largely reconfirmed, although in particular circumstances they may be challenged. After all, descriptions that are rephrased again and again tend to drift out of focus. There is every reason to suppose that scholarly accounts are just as likely to suffer from this phenomenon and neighborhood gossip. In the past there was no “original state” against which to compare. To illustrate this possibility I conclude with an example from a Hucbald treatise from the tenth century. My source (Figure 14) comes from *e-Codices*, the Virtual Manuscript Library of Switzerland (<http://www.e-codices.ch>). Hucbald has a one-line entry in many textbooks used on my side of the Atlantic. His singular importance is usually summed up as “the father of musical notation” or “the inventor of staff notation.” What Figure 14 shows is the leaf of Einsiedeln, Stiftsbibliothek, Codex 169 (468),³⁵ on which this “staff” is found.

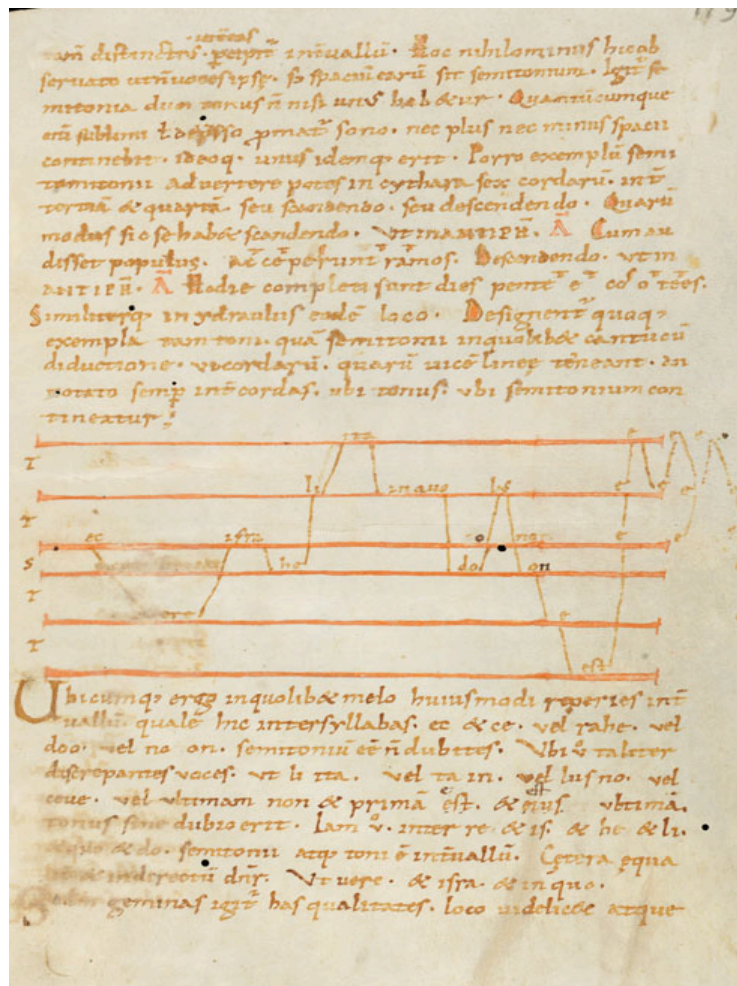


Figure 14. Leaf 119 of Einsiedeln, Stiftsbibliothek, Codex 169 (468), showing Hucbald’s exposition of “staff notation.” Reproduced under a Creative Commons 3.0 non-commercial license.

The topic under discussion is how to prepare novices to accompany chant on a harp. The moot point is whether the graphical images constitute “staff notation” or simply a harp tablature in which care has been taken to distinguish between whole tones (T) and semitones (S). What is

³⁵ Permanent link: <http://www.e-codices.unifr.ch/en/doubleview/sbe/0169/119>. A free tablet app for e-Codices has been released recently and is available via iTunes.

perhaps more interesting than the pitch representation is the effort to map duration through the width of the continuous line representing the melody. Our knowledge of early music has remained for more than two centuries dependent on the transcriptions of Martin Gerbert,³⁶ and as a group they are most impressive.

The “inventor of staff notation” claim is largely unchallenged, although some recent specialists tread lightly on this point. Yves Chartier, for example, says that Hucbald “suggests *placing the syllables of chants between the lines of a six-line staff*, where the intervals of tone and semitone between each line are indicated in the manner of a *clef*.”³⁷ The latitude with which those with direct access to an original source can interpret their meaning on their own points to a world in which students will be constructing knowledge as they see fit, with less and less intermediation of their pedagogical forebears. Digital resources are rapidly becoming the teachers of us all.

Eleanor Selfridge-Field

(Stanford University, Center for Computer Assisted Research in the Humanities)

Eleanor Selfridge-Field is Consulting Professor of Music at Stanford University and managing director of the Center for Computer Assisted Research in the Humanities (an affiliate of the Packard Humanities Institute). She has made important contributions to historical, cultural, and digital musicology. Her publications in digital musicology include the book *Beyond MIDI: A Handbook of Musical Codes* (1997) and fifteen issues of *Computing in Musicology* (1985-2008), as well as various journal articles in musical informatics, music cognition, and music as intellectual property. She has served on a series of standards panels including (currently) the Music Encoding Initiative.

Her historical work concentrates on Venetian subjects including a book-in-progress on Antonio Vivaldi. Her *New Chronology of Venetian Opera and Related Genres* (Stanford University Press, 2007) won the Modern Language Association Distinguished Bibliography Prize in 2008. Its companion volume, *Song and Season* (also Stanford, 2007), shows how conventional calendar time was manipulated to accommodate local cultural priorities and how cultural tradition prevailed over the scientific reform of clock time.

She holds degrees from two US universities (Columbia and Drew) and a D. Phil. from Oxford University.

esfield@stanford.edu

³⁶ Gerbert’s text transcription of Hucbald originates in *Scriptores ecclesiastici de musica sacra* (3 vols., St. Blasien, 1784), available through <http://books.google.com>. The work includes several other items that may also be derived from this Einsiedeln manuscript.

³⁷ “Hucbald,” *Grove Online*, accessed on 19 May 2012.