

Schriften des Instituts für Dokumentologie und Editorik — Band 13

Versioning Cultural Objects Digital Approaches

edited by

Roman Bleier, Sean M. Winslow

2019

BoD, Norderstedt

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Digitale Parallelfassung der gedruckten Publikation zur Archivierung im Kölner Universitäts-Publikations-Server (KUPS). Stand 18. Dezember 2019.

2019

Herstellung und Verlag: Books on Demand GmbH, Norderstedt

ISBN: 978-3-7504-2702-0

Einbandgestaltung: Julia Sorouri and Stefan Dumont; Coverbild gestaltet von Vinayak Das Gupta.

Satz: Roman Bleier, Sean M. Winslow und Lua \TeX

Versioning Cultural Objects through the Text-Encoding of Folk Songs

Richard Breen

Abstract

This paper will present and discuss experiences studying different versions of folk songs as cultural objects, and will investigate how using specific Digital Humanities tools may assist the versioning of intangible oral tradition. This was primarily achieved using The Versioning Machine, a framework and an interface for displaying multiple versions of text and audio encoded according to the Text Encoding Initiative (TEI) Guidelines. Through encoding a set number of songs in The Versioning Machine and displaying the results online, new questions and conclusions could be made to version cultural material with an emphasis on trying to trace the evolution of cultural ideas through subsequent iterations of ideas. Using examples from the project *Documenting Transmission: The Rake Cycle*¹, this paper will examine the effectiveness of using a specific existing versioning tool to model and map the differences between versions of folk songs and examine the intangible nature of performance and oral tradition. How do these digital versions change or reinforce our perception of a song cycle and transmission processes in general? This paper will give a broad overview of the Documenting Transmission project and some of the musicological and technical considerations that were made over the course of the project.

1 Introduction

The malleability with which the term *version* is used across disciplines within the humanities is certainly valid cause for careful evaluation to assert some form of working definition across them. The advent, and continuing growth, of inter-disciplinary scholarship forms new perspectives, but also informs the treatment of humanities data. Given the vast amount of different media through which humanities data may now present itself, a broad treatment of what the term *version*, and the subsequent process of *versioning*, is should be paramount. Ontologically, the attempt to digitally classify a group of related cultural artefacts as versions of one another is hindered

¹ A link to the project may be found in the bibliography. N.B: for best functionality, please view in Mozilla Firefox.

firstly by the unquantifiable nature of much humanities data. With cultural material, particularly that of oral tradition, authority is often lost through ill-defined chronology. Attempting to digitise this intangible cultural heritage is a vast undertaking, particularly with vast amounts of information with often-loose structure based on oral tradition. Despite the ontological woes of dealing with humanities data and its intrinsic differences to scientific data, the treatment of cultural material in a digital environment at least allows us to re-approach older concepts, asking old questions and finding new answers. In an ethnomusicological context, we may begin to digitise and map out tradition in new ways. Approaching versioning as a process is an understanding that the contextualisation of a cultural artefact is essential to the perception of the subsequent versions of that object. The modelling of these processes is therefore becoming an increasingly relevant issue among Digital Humanities scholars. The digital offers myriad new ways in which one may perceive artefacts as versions of an original artefact, and digital scholarly editions help embellish the narrative of an object's history and help form new perspectives for that history's treatment.

This paper will address versioning cultural objects through the case study of the *Documenting Transmission: The Rake Cycle* project, which text-audio encodes folk songs from the well-known "Rake Cycle" and visualises the transmission processes that occur between the songs in the *Versioning Machine* (Schreibman et al.). The term "Rake Cycle" refers to a nineteenth century English folk ballad entitled "The Unfortunate Rake," which through transmission processes has evolved into many variants in several sub-genres of music. While there are variants in languages other than English, this project text-encoded the lyrics and audio of thirteen English-language versions from Kenneth Goldstein's "The Unfortunate Rake" album, chosen for their motivic, musical, and geographical similarities and differences. The intention of this project is to highlight and display motivic and tropic information across the texts and music as opposed to simply the music itself. Previous and ongoing digital versioning projects, such as *The Thomas MacGreevey Archive* and *In Transition: Selected Works by the Baroness Elsa Von Freytag Loringhoven* (Clement) that have used text-encoding to mark-up versions of poems have dealt with the relationship between witnesses to one overarching textual idea. Tanya Clement, when referencing the *In Transition* project, describes the relationships between draft versions of poems as *Textual Performance Theory*, wherein the view is taken that the relationship and dialogue between each version of each poem is as important as the final product (Clement, "In Transition"). There is no definitive version of each poem, it is the differences between them that becomes interesting, whether it is the semantic differences between stanzas or instances of words, or the visual aspect of the words on the page themselves (Clement, "Knowledge Representation" 3). The *In Transition* project presents twelve unpublished poems written by Freytag-Loringhoven and marks them up according to the Text Encoding Initiative's P5 guidelines. The poems are then represented and edited critically

in the *Versioning Machine*, a customisable framework used to display text encoded according to the TEI guidelines (TEI Consortium), facilitating the close comparison of several texts, or *witnesses*, with a diplomatic or authoritative text (Schreibman et al., *Versioning Machine 5.0*, “Documentation”). Each poem contains several different *sketches* of that particular poem at different draft stages, the point of which is tracing the genesis of each poem from the earliest sketch to the completed poem. This assumes a process wherein the relationship and dialogue between each version of each poem is as important as the final product. Susan Schreibman, creator of the *Versioning Machine*, states that it “[e]nables [...] a theory of social-text editing wherein no version of the work carries more authority than another: each version of the work being a witness to a textual moment” (Schreibman, “Re-Envisioning” 93). The texts and sketches that are formative to a completed piece of work are therefore seen as a process that is as important as the result. This provides a complete textual history rather than establishing a definitive text. In marking-up literary work, a new digital version of said work is also created.

In classical music study, similar research has been made in *sketch studies*, wherein a composer’s draft manuscripts are observed relative to a completed work to form a genetic story of that work (Kerman 174). Some digital projects have begun the migration of this theory to the digital, marking up scores according to the Music Encoding Initiative (MEI) Guidelines. The *Beethoven Werkstatt* (Appel et al.) employs an MEI-based digital mark-up of the manuscripts of Ludwig van Beethoven. Similarly, the *Online Chopin Variorum Edition* (OCVE) marks up the many published versions of Frédéric Chopin’s scores to compare publication histories and the minute publication differences between the composer’s scores. Perry Roland explains that the true potential of the MEI and its implementation in the OCVE project is to “encode multiple versions of a musical work and generate multiple outputs” (Roland 8). On a functional level, this is a very similar principle to that of the aforementioned *In Transition* project; it is the addition and subtraction of material in Chopin’s differing printed editions that generates an interest in documenting these different editions of that composer’s work. The result is that it becomes much harder to call any one of Chopin’s printed scores as a *definitive* work.

Musically speaking, the *Documenting Transmission* project was concerned with the digitisation and versioning of oral material, primarily through the transcription and encoding of lyrics and audio tracks from an album of folk songs that display what is known as the *Folk Process*. As folk songs travel geographically or are shared culturally, they may adopt completely new musical characteristics while maintaining a core lyrical or thematic story. Norm Cohen states that:

All folk ballads are distinguished by a tendency of singers or composers to reuse stock phrases and even entire stanzas from older ballads [...] borrowing

phrases and stanzas from earlier works is no more plagiarism than it would be for a *literary* poet to hunt for words in a thesaurus...

One such issue with this project was that none of the resulting variations of the ballads of the “Rake Cycle” can be considered as the definitive version due to most of the changes in either music or lyrics being subject to the process of oral transmission. The focus is not on one author or composer, but rather many different musicians re-interpreting and building upon a source material to create their own version. Where *Documenting Transmission* differs is that where MEI-centred projects are concerned with the marking up of set musical notation, this project is concerned with the comparison of textual differences in lyrics, with the music serving more as an aural guide to the lyrics and cultural motifs. This is more demonstrative of the transmission process than that a formal analysis of the music itself, in that there is in most cases no formal or authoritative notation of the music in question. As such, the TEI Guidelines and Text Critical Apparatus tag set is used instead of the MEI Guidelines to encode the lyrical texts, and the *Versioning Machine* is used to display the results.

There are key points to consider when versioning such texts, namely that:

1. Folk songs are performances and are generally not transcribed lyrically or musically as in the Western Art Music tradition.
2. Attempting to map oral tradition presents a huge problem of authority between versions often due to a lack of publication or performance history.

While there are countless digital repositories dedicated to the digitisation and preservation of folk music and material, such as the *Vaughan Williams Memorial Library* (VWML) or the *Comhaltas Traditional Music Archive*, there is often a tendency to amass material and less of a focus on the digital modelling and study of the relationships between songs in the vein of a digital scholarly edition. This project was therefore intended to marry the methodologies of ethnomusicology and previous text-critical versioning projects such as those listed above. While it would be possible to adapt *The Versioning Machine* to display MEI, the sheer size of musical variation between many of the versions used in this study means that attention to minute musical change is eschewed in favour of demonstrating broader motivic changes in lyrical tradition.

The most recent release of VM contains a new text-audio linking feature, which had originally been developed by members of the Modernist Versions Project and is now a standard component of VM 5.0.² This feature facilitates parallel reading of a version of a text and at the same time listening to an audio version. In observing how other projects utilised the *Versioning Machine* to visually represent the versioning process, *Documenting Transmission* utilises and extends some of these features of VM

² See v-machine.org/documentation/#enc_audio.

5.0 to allow the comparison of different motivic, lyrical, and musical features of folk songs.

2 The folk process

The study of folksongs as an example of the tracing of cultural lineage between versions is a long-studied area of musicology and ethnomusicology, for at least over one hundred years. In 1907, the seminal English musicologist Cecil Sharp noted that folk music is communal in two forms, in that of its creation and in its representation of the thoughts of the community. In *English Folk Song: Some Conclusions*, he quotes F. M. Boehme, stating “[f]irst of all one man sings a song, and then others sing it after him, *changing what they do not like*” (Sharp 10). Perhaps the most famous of Sharp’s observations in this book is his identification of three factors that govern the transmission of folk songs musically and the forms these songs took depending on their context. These were:

1. Continuity: the linking of the past and present.
2. Variation: that the changes that occur rely solely on the creative tendencies or impulses of an individual or group.
3. Selection: that the community in question chooses which music it plays, which in turn decides the form(s) the music takes as it survives. (Sharp 17–31)

Sharp’s observations are an effort to identify how internal and external factors effect or determine the direction of the process of oral transmission, such as war, politics, and societal change. Musicologists throughout the mid to late 20th century, while being undecided as to how exactly to define folk music, agreed that Sharp’s three characteristics were intrinsic to the creation of folk material (Cowdery 808). Charles Seeger is credited with coining the term *folk process* to define the “...process by which cultural artefacts [sic] are changed, whether minutely or in significant amounts, to form new cultural products” (Washbourne 457). A group of folk songs with similar tunes or similar lyrics may therefore be identified as being a part of the same song family, wherein each performer identifies their own version of a song.³

The *Documenting Transmission* project centred mainly on trying to digitally map the folk process in action across selected songs from a song cycle to show the transmission and transformation of literary and artistic material from person to person in both an

³ David Atkinson, speaking of the genetic links between folk songs, states that “a particular song as taken down from a particular contributor is most usually said to constitute that person’s version of that particular song type, wherein *type* means the constant elements that “unify [...] what is recognizable as the ‘same’ song” (4). It should be stated then that this definition of version is based around the folkloristic interpretation of a version. Since this project is concerned with representing this definition of a version, as opposed to representing textual witnesses, this definition of version is used when describing individual songs.

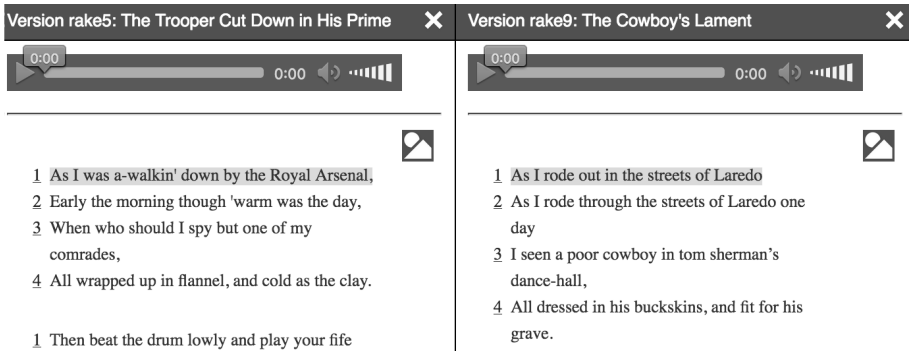


Figure 1: How songs, or Versions, are presented in the *Versioning Machine*. A user may press play on any line of text and hear the audio of that corresponds to that line of text. Similar lyrical ideas are highlighted across all versions corresponding to whichever track is selected.

oral and literary context. Essentially this means that as the stories, art, and musical traditions of a community are passed down from generation to generation, they are subject to organic changes in form, context, narrative and performance. *Documenting Transmission* is an investigation into displaying the interoperability of the narrative, musical, and visual functions that song lyrics serve. In some cases, certain elements may be completely changed or left out completely between versions or the music may be entirely different. In a digital context, this raised the question of whether existing DH tools could be used to reflect both musical and literary change in a way that extends our consideration of these processes. The *Versioning Machine* framework allows a user to deconstruct and observe similarities and differences between versions concurrently, in such a way where a user may be aware of, or disregard, their historical context at will. They may press play on any one of the versions and actively see where lyrical similarities or differences occur in real-time across every version (see figure 1).

The collection of songs used for this project comes from a cycle of which the earliest surviving version is a fragmented verse from a 1790 broadside pamphlet, entitled “My Jewel, My Joy” (Lodewick 98). The subsequent full, or more substantial, versions in the cycle have been the subject of analysis for quite some time, with major ethnomusicological criticism coming in the mid-twentieth century by Kenneth Goldstein and Kenneth Lodewick. The most popular of the early versions of the song dates from 1840, most commonly titled “The Unfortunate Rake” or “The Unfortunate Lad” (Harwood 26; Roud Broadside Index B130326). The versions that were encoded for the project are taken from Goldstein’s educational compilation album *The Unfortunate Rake* released in 1960 by Smithsonian Folkways, its intention being a “study in the evolution of a ballad”. This album’s liner notes provide much background information

about each version and provide transcriptions of the lyrics, as well as contextual information for the performances. Encoding these songs as versions therefore required the encoding of both the lyrics (content) and the audio (performance) in the form of MP3 files.

The traditional story of this homiletic ballad recounts a protagonist who happens upon a former comrade dying of venereal disease at the side of the road. Subsequent verses detail how the invalid came to be “disordered” by a camp follower, assumedly a prostitute (Goldstein, “Liner Notes” 2). Depending on where the ballad travels to, the role of the soldier changes; it travels to sea and becomes about a sailor, or to the Americas to become a cowboy or miner. Other functions of the lyrics such as places, people, or events also change depending on that version’s environment or performer. This lyrical story and its countless reinterpretations of the lyrics form certain tropes and possibly hundreds of various incarnations and parodies across the world over a period of at least two hundred years. This identification of tropes and motivic functions in ethnomusicological terms is rooted in anthropological and folkloristic study. In his seminal *Morphology of the Folk Tale*, the folklorist Vladimir Propp made efforts to break down the basic plot components of Russian folk tales to reduce them to their most basic functional state. He observed that fairy tales from Russia were made up of thousands of literary functions that interacted with one another in countless combinations throughout the canon of Russian folklore. This fluidity of actions as opposed to unchanging material applies directly to the folk process and oral traditions. Propp argues that actions are more important than the dramatis personae, this too can be said of the Rake Cycle (7). The invalid in any of the folk songs may be a soldier or a sailor, his gender may be reversed, or the setting of the story may completely change. What is important is that there are certain constants and variables that may be found in every version. These include a military funeral, an invalid, or a warning to the listener. The *Versioning Machine* is therefore well suited to show these changes, as many small minute links can be location-referenced in context to one another.

3 The *Versioning Machine* and the text-encoding process

The *Versioning Machine*’s primary function is to display text encoding to facilitate close comparison of several texts, or witnesses, with a diplomatic or authoritative text (Schreibman et al., *Versioning Machine 5.0*, “Documentation”). It does this by taking a TEI document and transforming it into a HTML page that can be opened in a browser. Some features include the ability to add annotations and notes to encoded text, as well as providing an image viewer that can pan and zoom. These features were utilised in the *Documenting Transmission* project, with further aspects being implemented such as a colour coded motif reference index. This basic visualisation was used to identify

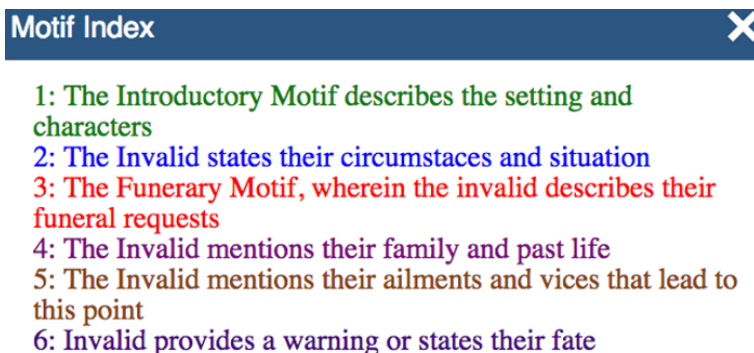


Figure 2: The Motif Index indicates which motifs are assigned which colour.

different motifs and thematic ideas across versions. This was achieved by customising the *Versioning Machine* by adding a simple plugin that places two buttons on the *Versioning Machine*'s interface. These toggle on and off a visualisation of the motifs across the song cycle by highlighting certain stanzas or lines in a pre-determined colour (see figures 2 and 3).

The Critical Apparatus tag set is designed to provide editors with a structured method of recording differences or variations between multiple witnesses of the same text.⁴ Using this tag set allows an editor to encode multiple versions of a text in a single document. A temporal alignment of text and audio is created using location based referencing, wherein the text of a transcription is encoded in the TEI critical apparatus structure. The audio files themselves are declared in the TEI `<witDetails>` element. Via its `@target`, an audio file can be linked to a particular text version.

```
<witDetail wit="#rake1" target="#rake1">
  <media mimeType="audio/mp3" url="audio/unfortunate-rake.mp3"/>
</witDetail>
```

The *Versioning Machine*'s XSLT stylesheet detects if an audio file was aligned with a text and includes an HTML audio element at the transformation stage into the VM interface. The visual alignment between different parts of the audio track and the transcribed text happens via a jQuery plugin. By clicking on a section of a text, the section is highlighted and the corresponding audio is played. Additionally, TEI `<timeline>` and `<when>` elements are embedded to provide a set of ordered points in time. Every `<when>` element represents a piece of text and is linked to a fragment of an audio track. In that way a temporal alignment of text and audio is achieved.

⁴ See v-machine.org/documentation/#method2.

The screenshot shows a web application interface for comparing two versions of a folk song. At the top, there are navigation tabs: 'Line numbers', 'Bibliographic panel', 'Notes panel', 'Visualise Motifs', and 'Motifs Index'. Below these are two panels, each with a title and a close button (X).

Version rake2: The Unfortunate Rake

Timeline: 0:03 to 2:59

1 As I was a-walking down by St. James's Hospitalⁿ,

2 I was a-walking down by there one day,

3 What should I spy but one of my comrades,

4 All wrapped up in flannel, though warm was the day.

1 I asked him what ailed him, I asked him what failed him,

2 I asked him the cause of all his complaint.

3 "It's all on account of some handsome young woman,

4 'Tis she that has caused me to weep and lament.

1 "And had she but told me before she disordered me,

Version rake3: The Bad Girl's Lament

Timeline: 0:00 to 0:00

1 As I walked down to St. James' Hospital,

2 St. James' Hospital early one day

3 I spied my only fairest daughter

4 Wrapped up in white linen as cold as the clay.

1 So beat your drums and play the life lowly,

2 And play the dead march as you carry me along;

3 Take me to the churchyard and lay the sod over me,

4 I am a young maid and I know I've done wrong.

1 Once in the street I used to look handsome;

2 Once in the street I used to dress gay;

3 First to the ale house, then to the dance hall

4 Then to the poor house and now to my grave.

Figure 3: Motifs highlighted across two versions.

```
<1 n="2">
  <app loc="locA1">
    <rdg wit="#rake3">
      <timeline unit="s">
        <when since="#t1_rake2_2" interval="4" xml:id="t1_rake2_3"/>
      </timeline> St. James' Hospital early one day
    </rdg>
  </app>
</1>
```

In the code example above, *locA1* refers to a specific point in time that will correspond with another set of lyrics identified as *locA1* in another version. This is essentially how two pieces of text are identified and highlighted at the same time. By clicking on a section of a text, the section is highlighted and the corresponding audio is played. A user may press play on any of the encoded versions and, while audio is played on that version, corresponding or similar lines of text are highlighted across all versions in real time until a song ends. Alternatively, a user may press on any line from any version to hear that version instantly (see figure 4). Allowing a user to break each song up and listen to any line of any version at a given time in any combination,

1 "Don't muffle your drums and play your fifes merrily,	1 "Oh, swing your rope slowly and ring your spurs lowly,
2 Play a quick march as you carry me along,	2 Play the dead march as you bear me along,
3 And fire your bright muskets all over my coffin,	3 "Take me to the green valley, there lay the sod o'er me,
4 Saying: There goes an unfortunate lad to his home."	4 'Cause I'm a poor cowboy and I know I've done wrong."

Figure 4: The Versioning Machine allows a user to see where similar motifs occur across different versions.

independent of the context. In this way, new spontaneous *songs* may be formed, albeit with a sense of chance. It can be said a user is actively *remixing* the song cycle, creating new cultural objects in the tradition of the song cycle, brief though they may be.⁵

The lyrics for each song were first analysed alongside one another for any immediate similarities by hand. The first level of analysis was to observe each set of lyrics and compare them to one another to find similarities or differences. These were informed largely by a number of texts that have analysed the song cycle, predominantly Goldstein's album liner notes and Lodewick's "The Unfortunate Rake and His Descendants." In the case of this song cycle, according to Goldstein ("Liner Notes" 1) and Lodewick (99), the earliest surviving documentation of the ballad is a single surviving verse titled "My Jewel, My Joy." This was encoded as the first version and serves as a starting point to trace motivic development across subsequent version. As this does not have an audio track on the album, there is no audio for this in the project. The first version encoded with audio is Track 1, "The Unfortunate Rake." Goldstein asserts that this version is "sufficiently close enough to the original ballad to warrant its use as a starting point for an examination of the whole family of related parodies and recensions" (2). As such, this project used "My Jewel, My Joy" and "The Unfortunate Rake" as a starting point from which changes can be observed across multiple versions of the ballad.

4 Observations

In figure 4, the left column displays a version of the ballad in its oldest recorded form "The Unfortunate Rake," whereas the version on the right, "Streets of Laredo," is recorded nearly a century later. The lyrics of the version on the left are more

⁵ The TEI document for this project is available on GitHub at github.com/rudgebreen/DocumentingTransmission--The-Rake-Cycle/blob/44211aa243c7973c956aab710b29580ebc9123c1/RakeCycle.

antiquated, depicting a military setting. The lyrics of the ballad on the right illustrate the relocation and recycling of the funerary motif to suit a cowboy context. Clearly some lyrical constructs have been carried over from the older song on the left and have been transmitted into that of the right, signalling that the action that is carried over in the cycle is more important than the song making historical or contextual sense.

In allowing a user the ability to click anywhere on any given version at a time, the *Versioning Machine* demonstrates a *deformance* of these encoded songs (Schreibman 99). Essentially, the performative meaning of a text is broken by way of taking the text out of its syntactical context. However, this project not only breaks the syntactical context of the lyrics visually. The entire tonality of the reading is changed through the changing of the music in tandem with the changing of the lyrics. This does not necessarily mean that the aural and visual readings are always broken together. There may be a line in one version that is completely the same syntactically as in another version, but the corresponding audio of each version is totally different. For example; both Version 3, “The Bad Girl’s Lament,” and Version 12, “The Wild Lumberjack,” contain the identical lyric: “And play the dead march as you carry me along.” However, the audio for Version 3 contains a voice accompanied by a guitar playing in a minor key with a definite pulse, whereas Version 12 contains one unaccompanied voice performing in a major key.

In terms of basic musical analysis, these are two completely different pieces of music regarding melody, tonality, and timbre. Observing just the lyrics by themselves creates no narrative context, it is the playing of the music that gives it some sort of context to the listener subconsciously. This is an immediate reaction as opposed to one that is informed by any previous textual information, and can be expressed without any need of the rest of the text. The melody that a singer chooses or the tonal quality of their voice can inform much to a listener about that line of text. The singer may be portraying conviction, sorrow, happiness, et cetera. Through this one lyric and using only the audio as the distinguishing feature between the two a further emphasis is placed on the performative aspect of the text and it adds to the sense of repurposing of the lyrics. If the folk process is the reshaping of cultural material into new cultural artefacts, then the performative nature of this project is as important to the result as the lyrics themselves. The encoding of audio alongside the text is essential in giving the user context both consciously and subconsciously.

The *Versioning Machine* allows the user to see the strong correlation between content (lyrics) and form (performance) in each of these songs. This in turn highlights the importance of influence between these songs and their performers. This has a huge amount to do with cultural transmission. As the ballad travels to new places, slang is re-used and misinterpreted to the point where new words and readings are created. One such example is “Bright Summer Morning,” recorded in the Virgin Islands. In

1 "And had she but told me before she disordered me,	1 Come Papa, come Mama, and sit you down by me,	me,
2 Had she but told me of it in time,	2 Come sit you down by me and pity my case;	2 O come dearest mother, and pity my crime,
3 I might have got pills and salts of white mercury, ⁿ	3 My poor head is aching, my sad heart is breaking,	3 For my poor heart is breakin', my poor head is bendin',
4 But now I'm cut down in the height of my prime.	4 My body's salivated and I'm bound to die. ⁿ	4 For I'm deep in salwat ⁿ , and surely I must die."

Figure 5: Each Version alludes to venereal disease in their lyrics, although the genders are reversed in the middle and far right versions.

this version, the over-arching story changes perspectives of the protagonist to that of a prostitute who has been wronged by a man. In this recording, the performer describes that her body is in “deep in salwatⁿ.” This is a corruption of the term “salivated,” which Goldstein describes as mercury poisoning in the album’s liner notes (4). Both versions appear to have come from “The Unfortunate Rake,” wherein the Rake complains that had the prostitute disclosed that she had a venereal disease before “disordering” him, he could have “[...] got pills of salts and white mercury.” Despite changing the perspective of the protagonist of the song, “One Morning in May” and “Bright Summer Morning” keep the source of the characters’ “disordering,” in doing so it re-shapes the lyrics to tell a new story and changes the listener or reader’s perspective in the process. This is contextualised by encoding these notes in the TEI, which the *Versioning Machine*’s interface displays (see figure 5).⁶ Corruptions of lyrics in this way demonstrate this process of change, the performer is unwittingly creating a new cultural artefact that the user may observe happening in real time.

While influence and tradition clearly play a massive role in how each version of the folk-song and its successive descendants are presented, a question arises as to the individuality of a version or performance. At what point does a performer’s version become recognised as its own song? Nearly every one of these versions of the story has a different melody or structure, and yet they are still related whether through a fragmented refrain or motif. Many of the similar verses or lyrical motifs may appear at many different stages of a performance. It is not fair to assume that every one of these performers knows the heritage of the ballad or where the story elements come from. To what extent then do these songs [remain in the same tradition and when do they form new musical heritage?

One such example is the final version encoded in this project, “Gambler’s Blues,” which marks a shift toward the jazz and blues tradition that would eventually become the jazz standard “St. James Infirmary Blues.” The lyrics to both versions contain only scant reference to the 18th century Irish ballad (Harwood; Lodewick). The

⁶ Goldstein asserts that the song probably made its way to the Virgin Islands by way of British colonisers in the 19th century, when the islands were under English control (“Liner Notes” 4).

ballad has been re-purposed to the point where it breaks away into a new musical tradition at the genesis of a new genre. As Lodewick states, the historical lineage of this particular version is blurred to the point where no real conclusions of relation may be drawn outside of basic motivic evidence (Lodewick). While this may be true, the *Versioning Machine* facilitates real-time comparison of this version of the song against the others in such a way that a user directly sees and hears these differences. A user may observe evidence of the influence of oral tradition on performers. By transcribing these lyrics into the digital format, the encoding process is a continuation of the process undertaken by Goldstein; in creating the *Unfortunate Rake* album, he transcribed the performances themselves, including the slang used by the performers, which he then contextualised critically. In encoding this process, the project is therefore creating a new cultural object.

While the *Versioning Machine* serves the purpose of comparing each of the sets of lyrics and audio from each song, this only displays the folk process at a textual and musical level through the thematic differences the lyrics represent, and the musical differences assumed by each performer. This facilitates close observation of the characteristic differences between versions on a micro level. Visualising the data is done in part by providing a colour coding of the motifs across the lyrics in the *Versioning Machine*. It was also important to try and use a tool that would map these changes at a macro level. This was done by mapping each of the versions used in the project in *StoryMapJS* to show the geographical distance over which the song cycle accrues these cultural motifs and tropes that define it. *StoryMapJS* is a free online tool that allows a user to tell stories on the web that highlight the locations of a series of events (Knight Lab). Through mapping where in the world each version of the song is iterated, new conclusions about the motivic links observed in the *Versioning Machine* may be drawn based on where the ballad travels to. A large element of the folk process is typified by where music travels to and the changes that can occur as music is transmitted through different cultures. *StoryMapJS* was useful in representing this, as it gives some locational indication of how far the ballad travels over the course of the versions used in this project. It also facilitated the implementation of images of the lyrics from each of the ballads and contextual information to be inserted alongside these for the user. As most of the ballads are either biographically situated by Goldstein in his liner notes or are directly in the lyrics themselves, this map was intended to give a broad indication of the song cycle's migratory span.

This is not an exhaustive representation of every version of the songs from this cycle. It serves as more of a contextual tool to embellish the section of the project that uses the *Versioning Machine* (see figures 6 and 7). The user is presented with bibliographic detail about each version of the ballad but also is provided with the lyrics above this information (see figure 7). Through the story map the user is brought on a curated geographical journey of the history of the song cycle to supplement the



Figure 6: StoryMap JS acts as a curated guide through the migration of the song cycle.

material in the *Versioning Machine*. What is apparent from this visualisation is that after the song-cycle makes its way to the US, there is no exact traceable direction that the ballad takes, outside of thematic differences, which can also be inferred from studying the cycle in the *Versioning Machine*. Examples of this are Versions 8 and 10, “St. James’ Hospital” and “The Streets of Loredó.” Despite both variants being recorded in the same state in the US—Texas—they are vastly different in terms of content. What this visualisation indicates is that no matter where it is, these motifs are travelling to at a given time, they are transmitted by the performers based on their own intent. This is demonstrative of the melting pot of cultural differences that typify a huge country such as the US, and is interesting to realise this visually outside of the *Versioning Machine*.

5 Conclusions

This project can be viewed as a prototype for musicological study of folk songs in a digital space. The most beneficial aspect of this is the provision of a unique way



Figure 7: StoryMap JS allows key points to be mapped to give a broad idea of distances travelled, as well as the inclusion of biographical information.

of observing motivic changes and representations through using TEI encoding and visualisation software. The motifs that carry over from performer to performer are indicative of the pervasiveness of certain themes or ideas that typify this song cycle, and viewing this process on a broader level gives an insight into how these processes function as part of a larger commentary on transmission processes.

One of the major aspects this project could expand on and address is the sheer amount of musical variation between versions. While technically both the music and lyrics of each version are encoded, the software itself limits how actual musical differences between versions are represented. These complications make it hard to digitally model the intrinsic link between the lyrics and the music outside a purely aural indication that is up to the listener. Some visual indication of the musical differences would be extremely beneficial in representing this aspect of folk song. While this does not affect the overall research question of the project, it does provide an avenue for later research into the representation of musical differences between folksongs to demonstrate musical change. This text-encoding method therefore suits

the research question in that illustrates motivic and tropic change across both text and music. However, the adaptation of the *Versioning Machine* to display different encoding guidelines such as MEI would certainly be an interesting area in terms of the close-reading of sheet music. Authority between versions is also an issue: in some cases, with certain versions of these folk songs very little is known regarding their date of composition or from which regional tradition each variation of the ballad is coming from after a certain point in a song's emigration. Other music-centred projects, such as the *Online Chopin Variorum Edition*, wherein the cultural material represented takes the form of published music, is aided through historical publication as well as dating the primary manuscripts themselves. In this way, definitive chronology of the material may be asserted. Mapping the data in this project does not lead a user to definitive historical conclusions, but it does help clarify how influence and tradition affect transmission processes.

The utilisation of location-based referencing caused some issues throughout the encoding process, not because of the functionality itself but because of the encoding of both audio and text. Location-referenced encoding in the *Versioning Machine* was easily the most time-consuming aspect of the encoding process. The reason for this is that because this project requires each individual line of text from each separate version to align with a specific place in each of the audio tracks, each line of text then needed to be encoded according to a specific time in the audio file between the previous and following lines of text. This is not an issue in purely text-based projects, as you can display how witnesses change across different lines of the text without the need of linking text to a defined section of audio, multiple lyrics that were similar could have been encoded in the same apparatus <app> element (Schreibman et al., *Versioning Machine 5.0*, "Documentation"). Some of the elements or tags that make up the tag set contradicted the research being done syntactically. For example, each version was encoded with the <witness> element, although this research asserts that these versions cannot be labelled as *witnesses*. This was worked around by changing how the versions were displayed in the actual interface itself by editing the HTML file, which allowed the *Witness List* to be changed to *Version List*, and each *Witness* could be changed to *Version* (see figure 8). However, in the TEI document itself the versions are still labelled as witnesses. This is indicative that in this case, extending the TEI ODD in the future to better reflect the mark-up of these versions in the *Versioning Machine*.

There are certainly benefits to exploring transmission processes through observing the inheritance across folk traditions in digital spaces. The pedagogical potential of creating digital editions in this vein would certainly help to give musicological study more presence in the Digital Humanities. While this can certainly be refined and developed in the future, it is indicative that the TEI can be used to identify and represent the links between versions that are not formally published or part of larger

editions. This project demonstrates that motivic elements of folk songs can become popularised to the point of becoming tropes, and that these tropes become more and more representative of a song cycle as it ages and travels. What is reinforced is that this type of multi-media based tool has worth within the humanities for explaining how different aspects of intangible concepts such as the folk process can be both contextualised and represented in a digital space.

Mapping out this song cycle in the *Versioning Machine* gives a greater sense of visual and aural contextualisation regarding these recordings as cultural objects. This method of modelling and treatment of sources is not restricted to any one case study or song cycle. This research demonstrates how creating resources that contextualise and visualise transmission processes can display the shaping of a musical or literary tradition, particularly in ways that apply to enthusiasts in musicology or folklore studies. In identifying changes and similarities on a small scale through location-based encoding in the *Versioning Machine*, a close reading of the texts and performances is facilitated. The creation of a visualisation through colour coding provides an aspect of distant reading, wherein the user can trace these motifs as larger groupings of narrative. Presenting the larger bibliographic narrative and migratory patterns of the song cycle in *StoryMapJS* allows the user to spatially contextualise it. Through using these different technologies in tandem, transmission processes can begin to be mapped within the digital spectrum.

Bibliography

- Appel, Bernhard R., et al. *Beethovens Werkstatt*. beethovens-werkstatt.de/. Accessed 12 Nov. 2017.
- Atkinson, David. "A Critical Edition of Folk Songs and Plays: Imaginings and Constraints." Presentation at the *Book History Research Network Study Day on Electronic Texts*, Institute for Textual Scholarship and Electronic Editing, University of Birmingham, 2006, pp. 1–13. www.abdn.ac.uk/elphinstone/documents/Birmingham.pdf. Accessed 12 Nov. 2017.
- Breen, Richard, editor. *Documenting Transmission: The Rake Cycle*, 2016. dhprojects.maynoothuniversity.ie/rbreen/Documenting%20Transmission:%20The%20Rake%20Cycle/samples/unfortunateRake.xml. Accessed 30 Mar. 2017 [Unfortunately the website is not working anymore. The source code is available on GitHub: github.com/rudgebreen/DocumentingTransmission-The-Rake-Cycle/. Accessed: 26 Apr. 2019].
- . *Documenting Transmission: Digitally Tracing and Displaying the Folk Process*. M.phil. thesis, Maynooth University, 2016. www.academia.edu/34643263/Documenting_Transmission_Digitally_Tracing_a. Accessed 26 Apr. 2019.
- Burns, Robert G.H. "Continuity, Variation, and Authenticity in the English Folk-Rock Movement." *Folk Music Journal*, vol. 9, no. 2, 2007, pp. 192–218. JSTOR, www.jstor.org/stable/4522807. Accessed 30 Mar. 2018.

- Clement, Tanya. *In Transition: Selected Poems by the Baroness Elsa von Freytag-Loringhoven*. 19 May 2016. digital.lib.umd.edu/transition/source. Accessed 30 Mar. 2017.
- . “Knowledge Representation and Digital Scholarly Editions in Theory and Practice.” *Journal of the Text Encoding Initiative*, issue 1, 2011. jtei.revues.org/203. Accessed 12 Nov. 2017.
- , editor. *In Transition: Selected Poems by the Baroness Elsa von Freytag-Loringhoven*. digital.lib.umd.edu/transition. Accessed 30 Mar. 2017.
- Cohen, N. “Folk music.” *Grove Music Online*, 2016. www.oxfordmusiconline.com/subscriber/article/grove/music/A2241135. Accessed 30 Mar. 2017.
- Comhaltas Traditional Music Archive*, 2000–2009. archive.comhaltas.ie. Accessed 21 Aug. 2019.
- Cowdery, James R. “Kategorie or Wertidee? The early years of the International Folk Music Council.” *Répertoire International de Littérature Musicale*. www.rilm.org/historiography/cowdery.pdf. Accessed 12 Nov. 2017.
- Goldstein, Kenneth S. *The Unfortunate Rake: A Study in the Evolution of a Ballad*, liner notes and musical recording, Smithsonian Folkways, 1960.
- . “Still More of “The Unfortunate Rake” and His Family.” *Western Folklore*, vol. 18, no. 1, 1959, pp. 35–8.
- Grier, James. “The Critical Editing of Music.” *The Critical Editing of Music: History, Method, and Practice*, edited by James Grier, Cambridge University Press, 1996, pp. 10–3.
- Hand, Wayland D. “The Cowboy’s Lament.” *Western Folklore*, vol. 17, no. 3, 1958, pp. 200–5. *JSTOR*, www.jstor.org/stable/1496046. Accessed 30 Mar. 2017.
- Hankinson, Andrew, et al. “The Music Encoding Initiative as a Document-Encoding Network.” *12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, Miami, 2011. www.mirlab.org/conference_papers/International_Conference/ISMIR%202011/papers/OS3-1.pdf. Accessed 12 Nov. 2017.
- Harwood, Robert W. *I Went Down to St. James Infirmary*. Harland Press, 2008.
- Kerman, Joseph. “Sketch Studies.” *19th-Century Music*, vol. 6, no. 2, 1982, pp. 174–80. *JSTOR*, www.jstor.org/stable/746277. Accessed 30 Mar. 2018.
- Knight Lab. *StoryMapJS*. storymap.knightlab.com/advanced. Accessed 30 Mar. 2017.
- Lodewick, Kenneth. “The Unfortunate Rake” and His Descendants. *Western Folklore*, vol. 14, no. 2, 1955, pp. 98–109. *JSTOR*, www.jstor.org/stable/1496993. Accessed 30 Mar. 2018.
- The Music Encoding Initiative (MEI)*. music-encoding.org. Accessed 30 Mar. 2017.
- MusicXML*, www.musicxml.com. Accessed 30 Mar. 2017.
- The Online Chopin Variorum Edition (OCVE)*. www.chopinonline.ac.uk/ocve. Accessed 30 Mar. 2017.
- Propp, V. *Morphology of the Folk Tale*. The American Folklore Society, 1968.
- Roland, P. “The Music Encoding Initiative (MEI) DTD and the OCVE.” *Online Chopin Variorum Edition*. 30 Dec. 2015, pilot.ocve.org.uk/redist/pdf/roland.pdf. Accessed 30 Mar. 2017.
- Vaughan Williams Memorial Library (VWML). *Roud Folk Song Index*. www.vwml.org/. Accessed 30 Mar. 2017.
- Schreibman, Susan. “Re-Envisioning Versioning: A Scholar’s Toolkit. *Digital philology and medieval texts*, edited by Arianna Ciula and Francesco Stella,” 2007. www.infotext.unisi.it/upload/DIGIMED06/book/schreibman.pdf. Accessed 30 Mar. 2017.

- Schreibman, Susan, et al. "The Versioning Machine." *Literary and Linguistic Computing*, vol. 18, no. 1, 2003, pp. 101–7.
- , et al. *Versioning Machine 5.0*. 2016. v-machine.org/. Accessed 30 Mar. 2017.
- , editor. *The Thomas MacGreevy Archive*. 2007. www.macgreevy.org/. Accessed 21 Aug. 2019.
- Sharp, Cecil J. *English Folk Song, Some Conclusions*. Simpkin & Co, London, 1907. archive.org/details/englishfolksongs00shar. Accessed 30 Mar. 2017.
- Sperberg-McQueen, C.M. "Textual Criticism and the Text Encoding Initiative." *MLA*, 1994. www.tei-c.org/Vault/XX/mla94.html. Accessed 30 Mar. 2017.
- TEI Consortium, editor. *TEI P5: Guidelines for Electronic Text Encoding and Interchange. Version 3.4.0*. 2018, www.tei-c.org/Vault/P5/3.4.0/doc/tei-p5-doc/en/html.
- TEI Roma: generating customizations for the TEI*. www.tei-c.org/Roma/. Accessed 30 Mar. 2017.
- Thompson, Stith. *Motif-Index of Folk-Literature*. Indiana University Press, Bloomington, 1955–58. sites.ualberta.ca/~urban/Projects/English/Motif_Index.htm. Accessed 30 Mar. 2017.
- Washbourne, Kelly. "Translation, the 'Folk Process', and Socially Committed Songs of the 1960s." *Mutatis Mutandis: Revista Latinoamericana de Traducción*, vol. 6, no. 2, 2013, pp. 455–76.