Folklore Tracks:
Historical GIS and Folklore Collection in 19th Century Denmark

Ida Storm [a,b,c,d]
UCLA
ida.storm@hotmail.com

Holly Nicol [c]
UCLA
nicol.holly@g.ucla.edu

Georgia Broughton [c]
UCLA
georgia.broughton@g.ucla.edu

Timothy R. Tangherlini [a,b,c,e]
UCLA
tango@humnet.ucla.edu

[a] Conceived of the project
[b] Developed methods and workflow
[c] Extracted and cleaned data
[d] Developed visualizations: maps, charts, graphs
[e] Wrote text

Keywords: historical GIS, folklore, history of folklore, ethnography, fieldwork

Abstract

The "golden age" of folklore collection in 19th century Scandinavia coincided with rapid changes in political, economic, and social organization as well as the rise of the Scandinavian countries broadly conceived of as "nations". The large folklore collections created during this period were a result of broad field collecting efforts across the region. Tracing the routes of folklorists as they conducted fieldwork helps us discern the developing conceptions of the nation and its cultural boundaries, as well as identify the areas that were most associated, in the minds of collectors, publishers, and scholars, with the cultural locus of the nation. Unraveling the fieldwork methods of early folklore collectors is not a trivial undertaking, and requires a combination of archival research and modern computational methods to reverse engineer the processes by which their collections were created. In this paper, we show how techniques from GIS used in conjunction with time-tested archival research methods can reveal how a folklore collection came into being. Our target corpus is

the folklore collections of the Danish school teacher, Evald Tang Kristensen (1843-1929) who, over the course of his fifty-year career, traveled nearly 70,000 kilometers, much of it on foot. We focus on determining how and where Tang Kristensen traveled in Denmark as he created his collection. We develop detailed route maps for 264 field trips, and project these onto appropriate historical base maps showing his movement through the countryside. This work considerably extends qualitative assessments of Tang Kristensen's collecting and provides a comprehensive view of Tang Kristensen's engagement with the fraught and changing borders of the Danish nation in the latter part of nineteenth century.

Introduction

Since the inception of the field of folklore in the late eighteenth century, folklore collections have played a significant role in the "imagining of the nation" (Anderson 1983: 196). Part of this can be attributed to the fact that these collections play an important part in mapping cultural boundaries and, as such, contribute to the essential process of determining the situated genealogy of a nation (Anderson: 1983: 173-5). In Scandinavia, the "golden age" of folklore collection of the nineteenth century coincided with rapid changes in political, economic, and social organization as well as the rise of the Scandinavian countries broadly conceived of as "nations". Svend Grundtvig's collections of Danish ballads (1853-1976), Peder Asbjørnsen and Jørgen Moe's collections of Norwegian folktales (1842-1852), Gunnar Olof Hyltén-Cavallius's collections of Swedish folklore (1844) and Elias Lönnrot's *Kalevala* (1849), to name but the most obvious candidates, all played central roles in the broad imagining of these nations. Importantly, each one of these collections was based, at least in part, on folklore fieldwork, where individual collectors, or networks of collectors, visited the countryside to document the narrative culture of the rural populations and to trace the boundaries of the nation as reflected by their fieldwork sites (Dundes 1977).

Folklore fieldwork collecting as a modern scholarly enterprise, separate from folklore anthologizing, had already emerged in the early decades of the nineteenth century. Inspiration for fieldwork in the Nordic region can be traced to the work of Wilhelm and Jacob Grimm in Germany. In the introduction to their first volume of *Deutsche Sagen* (1816), they extolled the benefits of fieldwork. Writing in critique of their own reliance on their anthologizing work on written sources, they opined that, "many [of the legends] would have been better and more precisely recorded from the mouths of the folk" (Grimm 1816, translated in Ward 1981: 11). Taking this observation as inspiration, a young bibliographic assistant at the Royal Library in Copenhagen, Just Matthias Thiele, who had already begun anthologizing Danish legends, decided to test that proposition. During a weekend retreat at

the manor farm of a friend in southern Sjælland, he took several long walking trips, collecting stories from the people he met as he visited the small houses and villages of the area. With this series of relatively short but successful field trips, he quite accidentally brought folklore fieldwork into the Nordic region.

Thiele's collection of Danish legends, *Prøver af danske folkesagn* (1817), based on these initial trips and several additional, equally abbreviated forays into the world of the agricultural working classes, resonated throughout Scandinavia, and inspired similar but more ambitious work across the region. Taking a cue from the Grimms' implicit suggestion that the rural populations were the more "authentic" guardians of national culture, folklore fieldwork in Scandinavia came to focus entirely on these populations (Hult 2003; Bendix 1997). In subsequent decades, and gaining particular momentum in the latter part of the century, folklore collectors such as the Norwegian, Moltke Moe, the Swede, Per Arvid Säve, and the Dane, Evald Tang Kristensen, turned their attention to increasingly systematic collecting (Österlund-Pötzsch 2014). In most of these cases, however, collecting took place in somewhat limited areas, thereby creating a rift between the regional scope of the fieldwork and the national scope of the publications and archives fed by those collections (Christiansen 2013: 117).

The outcome of the emergence of fieldwork as the primary method for the creation of folklore archives was, over the course of a century, twofold. First, it provided direct validation of nation-building projects, providing cultural confirmation of the physical mappings of the landscape that had begun in earnest in nineteenth century Scandinavia (cf Høje målebordsblade 1842-1899). Second, it led to the development of some of the world's largest folklore collections documenting the narrative culture and daily lives of thousands of rural people caught in the midst of the rapid social, economic and political changes of the long nineteenth century (1788-1901). Consequently, by tracing the routes of folklorists as they conducted fieldwork, it should be possible to discern the developing conceptions of the nation and its cultural boundaries, as well as to identify the areas that were most associated, in the minds of collectors, publishers, and scholars, with the cultural locus of the nation. As Anderson notes, this also fed the hunger for a means to infer the genealogy of the nation (Anderson 1983: 198-9).

As powerful as these collections were in the project of imagining the nation, they are not without their modern detractors. Indeed, several later folklorists have expressed skepticism about the utility of the collections as resources for the study of folklore, and see them more as artifacts of a historical period with little ethnographic value (Holbek 1979: 82). This skepticism is often based on perceived notions of how these collections came to be, rather

than a deep exploration of the actual practices of the collectors themselves, overlooking the fact that these collections are the product of the meetings between people (Christiansen 2013; Palmenfelt 1993). The goal here, then, is not to attempt a mapping of the collected folklore, a task that provided an initial impetus for the development of the earliest scientific method in folklore (Krohn 1926), but rather a mapping of a folklorist as he moved through the countryside documenting folk expressive culture (cf Tangherlini and Broadwell 2014).

Unraveling the fieldwork methods of early folklore collectors is not a trivial undertaking, and requires combining archival research with modern computational methods as a means for reverse engineering the processes by which the collection was created. In this paper, we show how techniques from GIS, used in conjunction with time-tested archival research methods can reveal how a folklore collection came into being, the limitations of the collection, and potential collection biases of the fieldworker. While these folklore collections are often presented as "national", thereby implying comprehensive coverage of an area that geographically defines "the nation" – and thereby serving the processes identified by Anderson – the methods we propose here can help delineate the actual scope of collecting practices, in some cases identifying the strong regional biases that inform the field practice of many folklore collectors. Conversely, given the "national" label that frequently accompanies published volumes based on these collections (to wit, *Danske sagn*, *Norske folkeeventyr*, etc.), the methods we devise can also help identify where cultural elites situated the locus of the nation (Tangherlini and Broadwell 2014).

Our target corpus is the folklore collections of the Danish school teacher, Evald Tang Kristensen (1843-1929) who, over the course of his fifty year career, traveled nearly 70,000 kilometers, much of it on foot, visiting approximately 4,500 named storytellers in 1,812 identifiable places, recording these stories in a little more than 24,000 field diary pages. In this work, we focus on determining how and where Tang Kristensen traveled in Denmark as he created his collection. We develop detailed route maps projected onto appropriate historical base maps showing his movement through the countryside. We add parish level population statistics to explore the extent to which population density and population change may have influenced his practices, and we incorporate transportation network data than can help us understand how changes in transportation technology, such as the emergence of a railway network, may have had an impact upon those practices. In all, we map 264 field trips, starting with his first recognizable trip in 1868 and ending with his last documented trip in 1916. This work considerably extends qualitative assessments of Tang Kristensen's collecting (Christiansen 2013; Tangherlini 2013a) and provides a comprehensive view of Tang Kristensen's engagement with the fraught and changing borders of the Danish nation

in the latter part of nineteenth century. We also believe that this work represents a key contribution to the development of the folklore macroscope (Tangherlini 2013b), adding an additional lens through which to view and interpret a folklore collection.

Methodology

Discovering field work routes: Data extraction

When we began this work, there was no existing catalog of Tang Kristensen's field collecting routes – we had to devise this by coordinating annotations in his hand-written field diaries with his four-volume memoir, *Minder og Oplevelser* (Kristensen 1923-1927). The work was based on the underlying notion that the majority of his fieldwork took place away from his home; in other words, he went to meet storytellers where they lived, worked or studied, and did not have them come to him. Consequently, we could represent his collecting practice as one based on a group of individual field trips that occurred not only in space but also in time. The primary resource, his memoir, is based largely on letters he wrote home detailing all his collecting trips, and includes the people he visited, the towns and villages in which he stopped, and the various people he consulted to find informants. It also includes information on the means of transportation he used to get around as well as the general dates of travel. Much of this data is "noisy" insofar as the descriptions are incomplete and filled with ambiguities, and the two main data resources, the memoirs and field diaries, provide conflicting information. Adding his personal correspondence to the mix often confuses rather than clarifies the route descriptions.

Our team began by devising "proto-routes" for each of the fieldtrips that we were able to identify in the memoirs. We tagged trip start and end dates manually, as well as all stops and the stop order for each trip. Automated methods based on named entity recognition algorithms, when applied to Danish, have an accuracy rate low enough that there is no appreciable difference for a corpus as small as the memoirs (1670 pages) between automatic tagging with manual correction and manual tagging. As we read through the memoirs, we inserted trip start and trip end tags. Once we had a series of candidate field trips, we then tagged each of the individual trips for people visited, stops, ordering of stops, and easily identifiable means of transportation. Not all of the people that Tang Kristensen visited told stories: in later work, we will explore his reliance on a social network comprised largely of schoolteachers, on whom he relied to meet local storytellers. It is worth noting that not all stops are associated with informants; at the same time, since he frequently met more than one person at a stop, single stops can be associated with multiple informants.

Occasionally, we discovered "nested" trips within the identified fieldtrips, where Tang Kristensen described related or additional trips within the context of larger trips: these "nested" trips were broken out and treated as separate field trips.

The result of this tagging work is a corpus of field trip descriptions and, after straightforward tagged text extraction, a simple list of start and end dates, places, people, and conveyances. A sample of this work is detailed below using tags to identify the place and person names:

> <start> En anden udflugt, jeg gjorde i sommerferien, gjaldt <person>Pastor H. Dahl</person> i <place>Skorup</place>, som jeg havde hørt omtale for hans arbejde med at udrydde fremmede ord af sproget. Jeg traf ham ogsaa, og han tog lidt koldt paa mig i førstningen, men tøede snart op. Efter at jeg havde overnattet i præstegaarden, <transport>kjørte</transport> jeg med præsten til <place>Tvilum</place> og var der i kirke. Degnen i sognet boede ovenfor bakkerne, og saa holdt vi uden for hans bopæl for at tage ham med. Det mærkede jeg, han var vant til. <place>Tvilum kirke</place> ligger saa ensomt der nede paa det flade land ved <place>Gudenaa</place>, og der var rigtig nok øde ikke alene omkring kirken, men ogsaa hele vejen hen ad <place>Sminge</place> til. Det var ogsaa en meget sandet og daarlig vej, saa det rigtignok var besværligt og kjedeligt for præsten at skulle færdes der, hver gang han skulde til <place>Vole</place>. Derefter kjørte vi til <place>Vole</place>, hvor han holdt eftermiddags-tjeneste, men jeg var imidlertid inde hos degnens. Derfra tog jeg ud til <place>Silkeborg</place> og saa ud til <place>Sinding</place> for at besøge <person>Kristoffer Munch</person>, som var lærer der, og som jeg havde gaaet paa seminariet med. Han havde udeladt sig med, at der var en <person>gammel kone</person> i <place>Øster-Bording</place>, som var god til at synge, og hende vilde jeg da have undersøgt. Hun gik omkring og spandt for folk, og jeg traf hende godt nok og fik mange gode viser af hende, saa det var et udmærket fund, jeg der gjorde. Jeg var jo ogsaa saa heldig senere at faa et billede af hende, for hun blev meget gammel, og endda var hun højt op i aarene, da jeg traf hende i 1881. Jeg var ogsaa inde hos <person>R. Geltzer</person>, som havde faaet oprettet en højskole i <place>Sinding</place>, og han sang en vise for mig, men hans

højskole vilde ikke gaa. Ogsaa lærer <person>Okkels</person> var jeg hos i <place>Øster-Bording</place>. Han var ellers ude fra <place>Gjellerup</place>. Hos <person>Munch</person> havde jeg et godt tilflugtssted, mens jeg var der, og har siden besøgt ham, hver gang jeg kom til <place>Ringe</place>, hvor han snart efter fik et rigtig godt embede og blev, til han omsider tog sin afsked. <end>

Extracting the tagged information results in an abbreviated list of people, places, and means of transportation:

<start>
<person>Pastor H. Dahl</person>
<place>Skorup</place>
<transport>kjørte</transport>
<place>Tvilum</place>
<place>Tvilum kirke</place>
<place>Gudenaa</place>
<place>Sminge</place>
<place>Vole</place>
<place>Vole</place>
<place>Silkeborg</place>
<place>Sinding</place>
<person>Kristoffer Munch</person>
<person>gammel kone</person>
<place>Øster-Bording</place>
<person>R. Geltzer</person>
<place>Sinding</place>
<person>Okkels</person>
<place>Øster-Bording</place>
<place>Gjellerup</place>
<person>Munch</person>
<place>Ringe</place>
<end>

This tagging and extraction was done for each of the identified fieldtrips and formed the basis for the proto-route corpus.

The proto-routes were then aligned with the field diaries, the notes that Tang Kristensen took while speaking with his informants and additional stops; in some cases, additional proto-routes were generated on the basis of this alignment. The proto-route information was then entered into simple tables with a series of ten proto-route features. The features were converted into column headings in accordance with the rules of field headings in ArcMap, which allows for a maximum of ten characters in each heading (fig. 1).

Place	Date	Transport	MO_Start	MO_End	Notes	Sequence	Db_start	Db_end	FT
Hadsten	9/18/1896		4_107	4_113		1			111
Mundelstrup By		Train				2			111
Borum					Stayed the night	3			111
Mundelstrup By						4			111
Aarhus						5			111
Testrup	11/20/1896					6			111
Beder						7			111
Malling						8			111
Fruering	11/22/1896					9			111
Satrup		Walk				10			111

Figure 1: Excel table with "proto-route"

To ensure that our group was working with synchronized files, we did the majority of the data entry and updating in Google Sheets; these sheets were then exported and saved as MS Excel worksheets, a data format supported by ESRI's ArcMap (the GIS platform with which we work). As seen in Figure 1, there are many stops that have no accompanying information, a reflection of the inconsistency of Tang Kristensen's notes describing his fieldtrips. In later work, we plan to align field stops with the digitized catalog of informants, which will allow us to fill in some of the gaps, as we can then make a more accurate alignment of the start and end points with pages in the field diaries. Since individual pages of the field diaries are aligned with individuals who are, in turn, assigned a place of residence on their catalog cards, this work will increase the accuracy of the final routes. Unfortunately, this work is not yet complete, given the complex nature of these informant index cards (Figure 2).
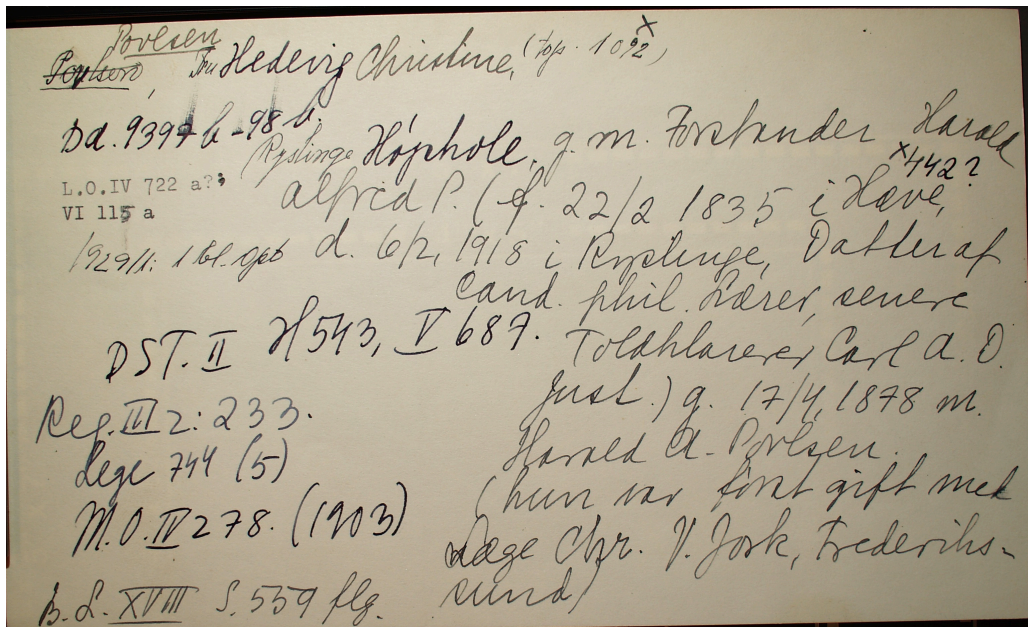
Figure 2: A sample informant card for Hedevig Christine Poulsen from Ryslinge, who is mentioned in the Memoirs, vol. 4, p. 278, and from whom Tang Kristensen collected during fieldtrip 160 in May 1903. From the index of informants in the Tang Kristensen collection.

Finding the locations for the stops that we extracted in our "proto-routes" was a significant challenge. As with most historical data, places can be difficult to locate: some are very small, names have changed, and some places have disappeared. Contemporary gazetteers are inadequate for the task and often confound rather than solve queries. To address this problem, we downloaded the historical place name database developed by the Afdeling for Navneforskning at Københavns Universitet, as part of the DigDag project, and used it to generate a customized "Address Locator" for nineteenth-century Denmark. All the stops of an individual fieldtrip could then be matched to the address locator in a single pass, generating a "best guess" solution for the proto-routes. The address locator's spelling sensitivity was changed to be more liberal during the matching process; this relaxation allowed us to locate names that, for various reasons such as a misalignment of orthography (e.g. å vs. aa), did not return a perfect match. Place ambiguity (multiple places matching a single record) was resolved through an "Interactive Rematch" interface, where the correct option was manually selected. To make these types of selections, each proto-route was inspected and individually compared to the fieldtrip notes. Impossible solutions were rejected and, when necessary, we closely examined the historical maps in concert with Tang Kristensen's descriptions from the memoirs, his field diaries and his correspondence to solve the ambiguity.

A simple aggregation of these geo-located places allowed us to rapidly check the general accuracy of the proposed solutions (Figure 3). Interestingly, this aggregation also confirmed that Tang Kristensen did the majority of his collecting in Jutland, as the islands are much more sparsely represented.
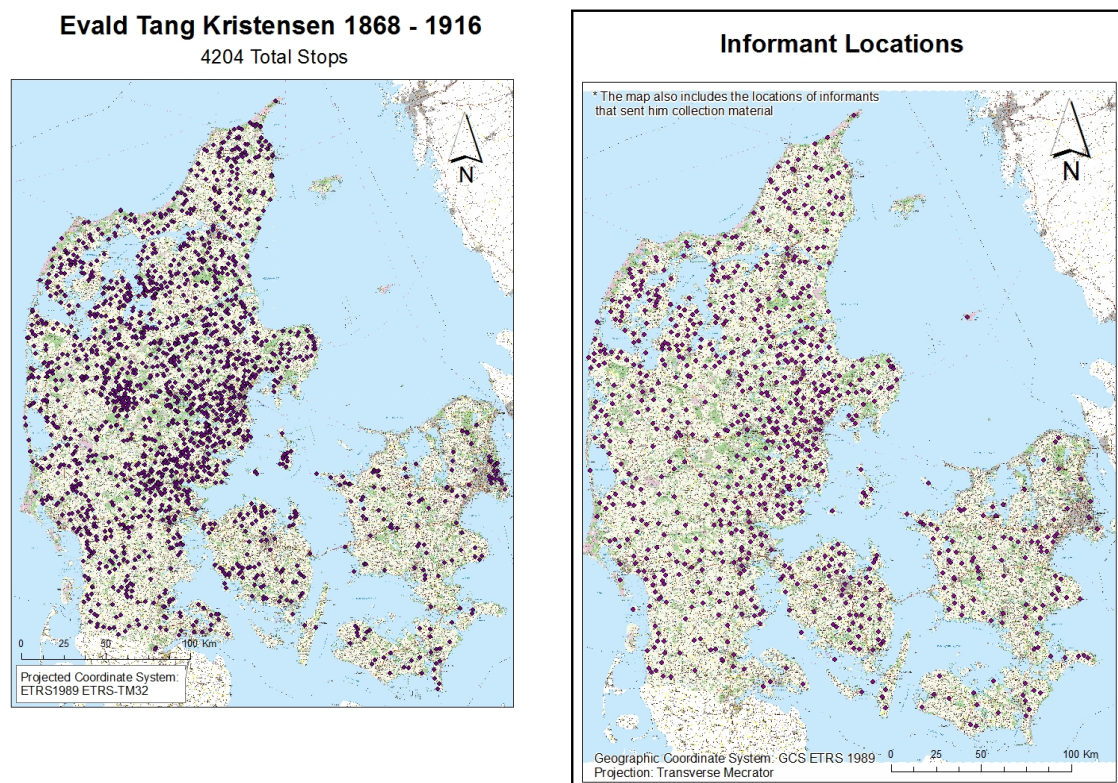

Figure 3: On the left, the aggregated stops for all of Tang Kristensen's travels, geocoded with the address locator. On the right, the 1,812 identifiable residence locations of his informants.

The scope of Tang Kristensen's collecting is immediately apparent in this visualization. When compared with a visualization of the place of residence of all his identifiable informants, it also reveals the degree to which Tang Kristensen moved between the countryside where he lived and collected and larger cities, including his numerous visits to Copenhagen. Future work with the informant index will allow us to identify additional informant locations.

From Proto-Routes to Routes

With the stops in a provisional sequential order, we created estimates of the "most likely route" for each trip. A basic assumption was that, unless otherwise specified, Tang Kristensen would take the shortest path between two points, an assumption that aligns with the underlying Dijkstra shortest-path algorithm in ArcMap (Dijkstra 1959). We used a transportation network from OpenStreetMaps, which we pruned by using the cadastral

survey maps of 1842-1899, the highest-resolution historical maps from the era, as a guide. Since Tang Kristensen occasionally traveled by boat, ferry lines based on ferry schedules and close study of historical maps were also added. By feeding the provisional sequential stops to the network analyst, we were able to create the most likely routes for each trip as a single line record. These routes were then visualized as a line with sequentially numbered stops (Figure 4).

Figure 4: Example of a route, here Fieldtrip 190, from November 1907, described in MO IV 350-7.

Simple descriptive statistics such as route length, as well as descriptors from our database such as dates of collection, field diary pages, and modes of transportation augment the visualization. These routes can be considered individually, in various groupings such as point of origin (usually Tang Kristensen's place of residence) or primary mode of transportation, or in the aggregate. Once the static routes are generated, they also provide a

new method for accessing the underlying collection and for classifying the stories in the collection.

Dynamic representations of fieldwork

Animations provide a dynamic representation of Tang Kristensen's movement through the countryside, capturing an aspect of his collection practice that static route maps cannot capture. Animations reveal, for example, the numerous times when he backtracked. To allow for sequential animations of data, a date field is normally used. However, as noted above, there are big gaps in the date field. To work around this missing data problem, we devised an additional "absolute order" field with the first stop assigned "1" and the last "4204" (in our work, we have identified 4204 unique stops where a unique stop is considered a visit on a particular day, irrespective of whether we have precise date information for that day). Before populating the absolute order field, the attribute table was sorted by fieldtrip and then by sequence using the "advanced sorting" function in ArcGIS. With the data in order, a simple Python script was written to fill the absolute sequence field. The routes were then extracted into inter-stop segments by starting an edit session for each individual route, with segments defined as the shortest path between any two stops. The corresponding absolute sequence value was added as the process proceeded.

In ArcGIS, the "Tracking Analyst" toolset includes an interface that translates shapefiles into temporal data. This data can, in turn, be used to construct animations. For this transformation, the toolset relies on a date field with certain data format requirements: the one that best matched our absolute sequence field was the year format, "YYYY", and so we aligned the date field with our absolute sequence field. Although the absolute sequence of stops numbered 1-999 did not comply with the data field requirements, we used a simple arithmetic work around to make our absolute sequence field conform with the date field requirements (http://wccas.org/etk_maps/route_animation.m4v).

Understanding fieldwork practice

Although it is well known that Tang Kristensen collected folklore throughout his life, and wandered widely through Jutland and some of the islands, there was no quantification of this fieldwork, nor was there any way to compare fieldwork during different periods in his life, during different seasons, in different regions, or as technology changed at the end of the

nineteenth century. In our work, we have identified and mapped 264 individual field trips. Because of the challenges of the data resources, particularly the discrepancy between Tang Kristensen's memoirs and his field diaries, there are likely to be several shorter trips that we were unable to identify, along with numerous interactions with people that did not meet the standard of a field trip. Similarly, there are probably some stops that are slightly out of order, and other stops that have either been inadvertently missed or added. Nevertheless, this nearly comprehensive overview of his field collecting provides a useful window onto his field collecting practice and allows us to interrogate his shifting ideas of folklore and the locus of culture at a very fine-grained level. Importantly, our work confirms Tang Kristensen's long standing engagement with Jutland, as opposed to Denmark as a whole, while at the same time revealing his changing evaluation of different parts of the country, if frequency and duration of visits can be taken as a proxy for evaluation. Always the pragmatist, Tang Kristensen was eager to get the biggest return on his investment of time. Consequently, he tended to go where people lived and adapted his fieldwork practice to allow him to cover the most ground using the emerging technologies of transportation, such as the train, as they were developed over the course of his collecting career.

Travel statistics

While Tang Kristensen's ability to walk long distances and persevere even in the worst weather is well documented in the folklore literature, our approach provides a clear method to quantify each of his trips and to quantify groups of trips. In ArcMap, route lengths can be calculated in the attribute table using "Calculate Geometry" and simple statistics of that field can be generated. The aggregated route shapefile provides an overview of all the travels he undertook (Figure 5):

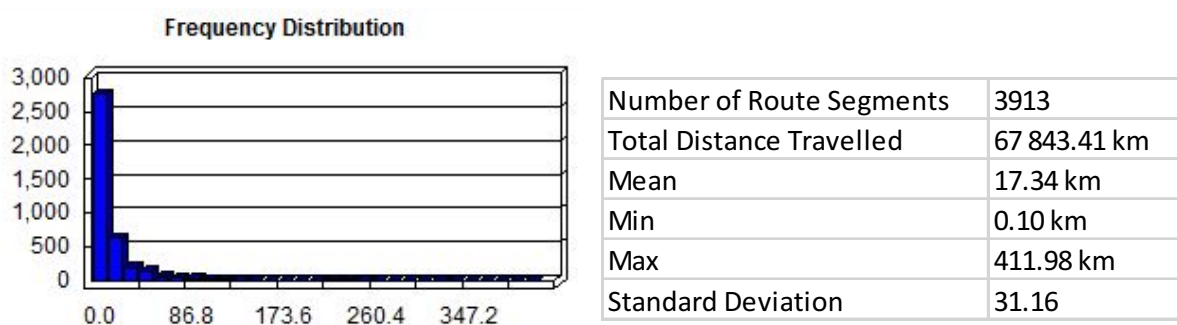| Number of Route Segments | 3913 |
Total Distance Travelled	67 843.41 km
Mean	17.34 km
Min	0.10 km
Max	411.98 km
Standard Deviation	31.16

Figure 5: travel statistics, number of route-segments on the y-axis and distance in kilometer on the x-axis.

Whereas the length of his inter-stop segments is highly variable, with a standard deviation of ~31km, the majority of his travel segments are short, with nearly half (~49%) being under ten kilometers. Indeed, the distribution of segments appears to follow a classic long-tail distribution. This observation of segment length aligns well with his memoirs, where walking is the predominant mode of transportation. Given the quality of roads and paths, the challenges of weather, and the actual time it takes to collect folklore, a distance of 10 km several times a day is probably at the upper limits of even a fit individual. In our calculations, we add additional options for modes of transportation. In future work, we plan to develop more detailed statistics regarding segment length and speed of travel based on travel mode. Currently, these calculations are limited by the lack of transportation data for many of the route segments. By implementing walking distance thresholds, however, we have been able to apply a likely bifurcation on certain segments for the classification: walking / not walking. Extrapolating from his descriptions of his longest walking trips as recounted in his memoirs, we have set the classification threshold to twenty kilometers.

With walking ruled out for the undefined route segments, we then model two additional overland transportation possibilities: train or wagon. We assume that he traveled by wagon on segments above our walking threshold, but we further surmise that the train gradually supplanted wagons as the rail network was built out. In a previous project, we developed a dynamic data set for the Danish train system, showing the distribution of train stations, the active railways and the year the railways and stations were opened. Clearly, travel over water required a boat and, for longer segments, was probably accomplished by steamer. Consequently, we classify field trip segments according to six categories: walk (definite), walk (20km threshold), boat/ferry, train, horse wagon and, for one segment, bicycle (in *Minder og Oplevelser* (IV:202), Tang Kristensen describes taking a bicycle from Marebæk to Skelby, and notes that this trip was the one and only time he used a bicycle in his life: see Figure 6). In future work, it may be possible to incorporate multiple modes of travel for particular travel segments, as well as to calculate travel time based on average speeds for each conveyance.
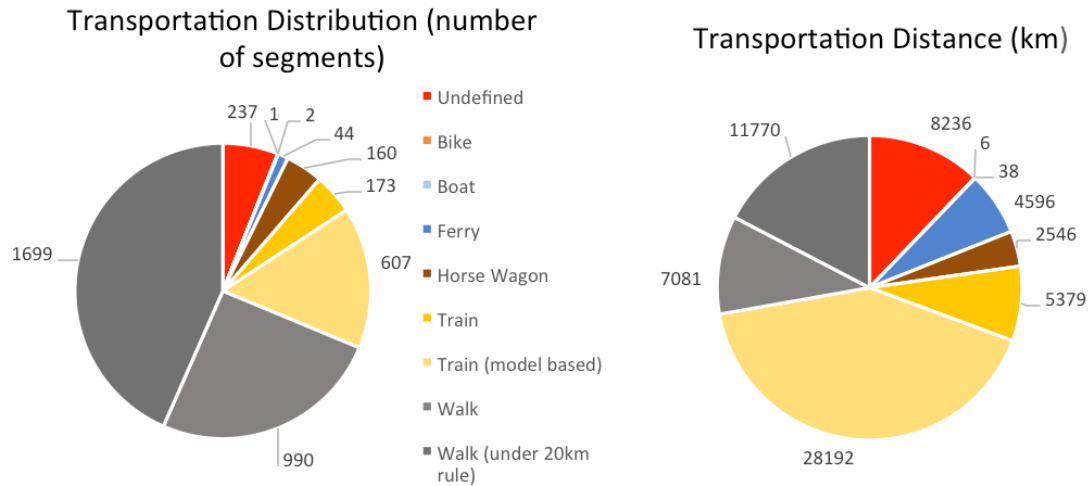
Figure 6: Transportation overview of Tang Kristensen's ~68,000 km traveled. On the left, number of segments by mode of transportation. On the right, total kilometers traveled by mode of transportation.

The impact of transportation technology and infrastructure development on folklore collecting is understudied. As we worked with these maps, we discovered that the overlap between train stations and undefined route segments is striking. As more train stations opened during the years of his travel, more and more of his segments included places that had trains. By intersecting undefined segment start and stop points with open rail lines and stations, we can solve those undefined segments that have a high likelihood of being traveled by train and reassign those segments to the train class. As a result of this realignment, we surmise that he increasingly used trains to get around, and that he did so far more than he mentions in his memoirs (Figure 7). Interestingly, this discovery means that the development of the railways, an important development in the context of the shift toward industrial production, urbanization and increasing contact with Europe, paradoxically also provided a greater scope for documenting the expressive culture of the rural populations that were seen by the intellectual class as threatened by those same developments.
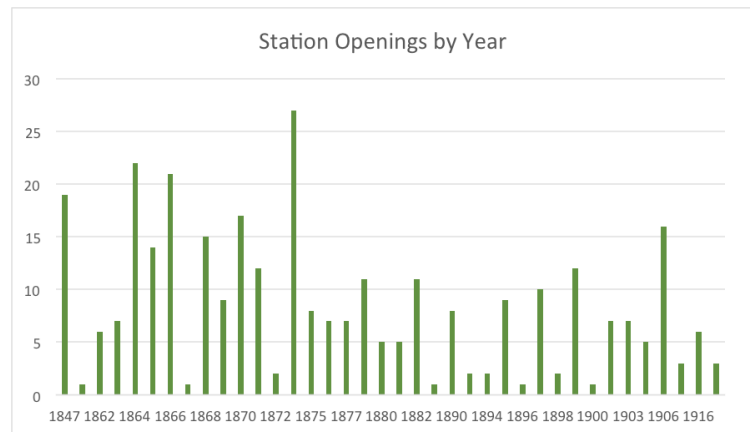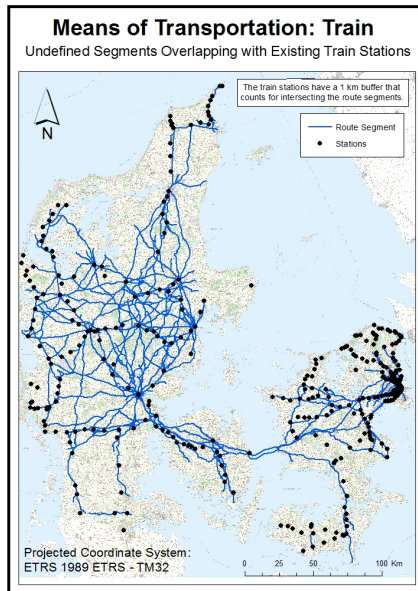
Figure 7: On the left, the undefined route segments that overlap with open train stations at the start and end point (blue). On the right, a histogram showing the distribution of train stations opened by year, ending in 1920. The number of stations is represented on the y-axis and the year is on the x-axis.

Seasonal and yearly variation in Tang Kristensen's field collecting

An intriguing outcome of our approach is the identification of temporal patterns in Tang Kristensen's fieldwork practice. These patterns can be characterized as micro-patterns related to the seasons and macro-patterns related to aspects of Tang Kristensen's life, including employment, family considerations and the inconsistency of external funding. Tang Kristensen recognizes in his memoirs that certain years were more productive than others in the context of his field collecting, but it is not immediately apparent from his discussions how to recognize these periods or how different they were in regards to fieldwork practice. To address these macro-questions of fieldwork practice periods, we binned the fieldtrips by year and then calculated two metrics: total distance traveled per year and total number of stops per year. The best way to get a grasp on this is by plotting the combined length of the route segments against a time axis (Figure 8). Since the distance is not necessarily correlated with how much collection Tang Kristensen did, it is also a good idea to look at the number of stops he did every year (Figure 9), which is more closely correlated to the number of informants with whom he interacted.
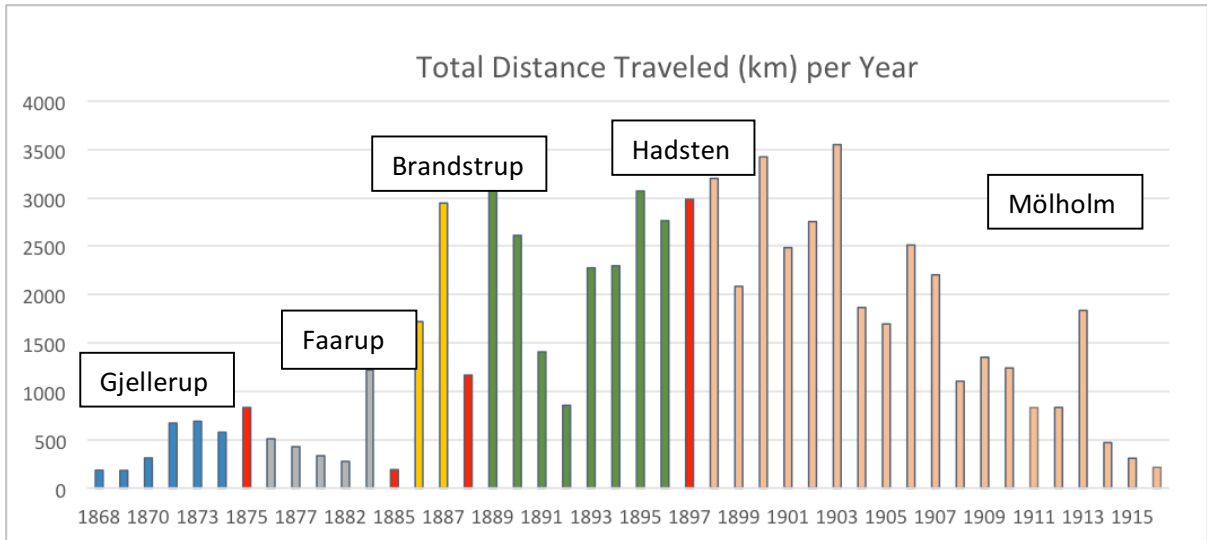
Figure 8: Distance traveled by year. There were no field collecting trips in 1872, 1878-1880 and 1884. We have added labels for Tang Kristensen's place of residence to the graph and have colored the bars according to place of residence. Red bars indicate overlapping places of residence.
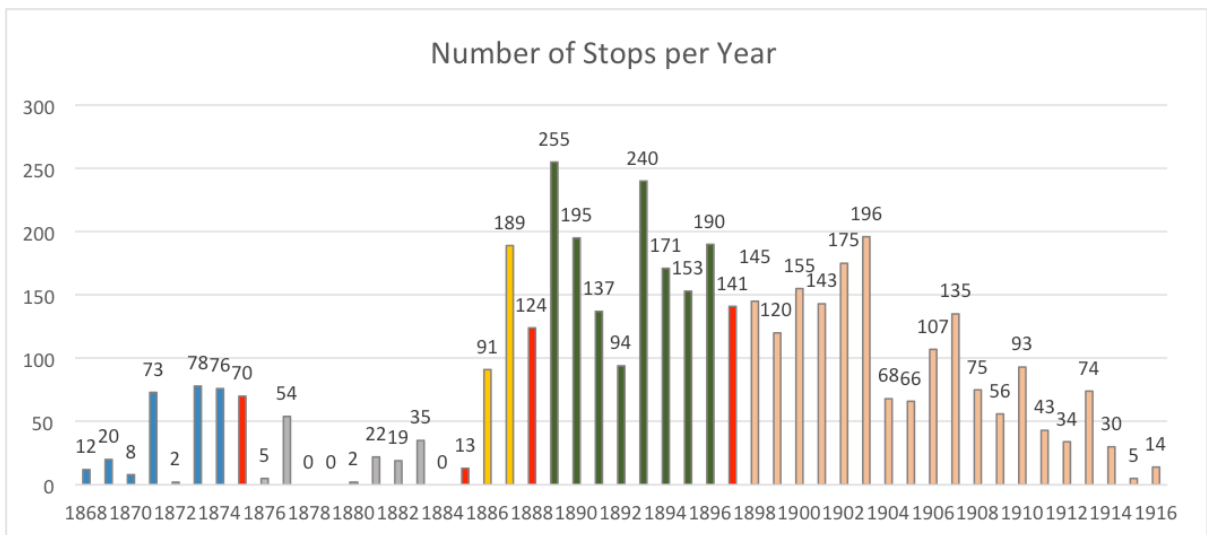

Figure 9: Number of stops by year. Places of residence are color-coded as above.

In the early years of his work (1867-1882), characterized by his early teaching career and residence in Gjellerup (until 1875) and later Faarup (1875-1885), a great deal of his collecting was in the service of Svend Grundtvig's efforts towards creating a comprehensive collection of Danish ballads. Interestingly, although he was developing a reputation as an excellent fieldworker, his travels were all relatively close to home; and the number of stops, while impressive, are markedly lower than later collecting periods. Tang Kristensen himself has noted the low point in his collecting career from 1878-1883, a phenomenon that can be seen in stark relief through these simple graphs. Indeed, his life in Faarup was marked by family obligations and a growing engagement in civic affairs, leaving little time or money for

the pursuit of folklore collecting (Rockwell 1981: 156-162). A growing strain between Tang Kristensen and Grundtvig also had significant impact on his collecting endeavors (Rockwell 1981: 160).

After Grundtvig's death in 1883, Tang Kristensen appears to have taken a pause for several years to reevaluate his folklore collecting but then hit the ground running, perhaps buoyed by his increasing ability to hire substitute teachers and the reduction in his involvement in local community politics that came with his move to Brandstrup (1883-1888) and later Hadsten Station (1888-1897). The move to Hadsten Station, in particular, seems to have led to an explosion not only in the number of field trips that he took, but also the length and the number of stops he made on these trips. It is worth noting that his move to towns that are very close to the rapidly developing railway infrastructure led to a marked increase in the number of trips and the aggregate annual distance traveled, a phenomenon that is most marked in relation to his move to Hadsten Station, a town on the Århus-Randers line which opened in 1862 and which provided Tang Kristensen a rapid line to northern Jutland. This proximity to the transportation infrastructure of late nineteenth-century Denmark continued with his move to Mølholm, just south of Vejle, three kilometers from Vejle station.

His fieldwork practice reveals an intriguing seasonal variability. Not surprisingly, there were several seasonal considerations that could have had an effect on Tang Kristensen's collecting. While the late spring and summer months would be ideal for field collecting from the perspective of weather and light, these were also months during which many likely informants were busiest with farm work. Consequently, while he could move easily and probably cover long distances during the summers, he may not have been able to find informants willing to spend the necessary time for his type of comprehensive repertoire collecting. Conversely, during the fall and early spring, Tang Kristensen was limited in his travels by his obligations to his teaching. Until he was able to secure significant funds to hire substitutes, he was limited to local fieldtrips. In the winter, when many people were indoors and thus likely to be available to tell stories and talk about their daily lives, the weather and the lack of daylight would be limiting factors. Given the conflict between these potentially limiting factors, it is hard to make any clear inference about seasonal productivity.

To address whether there were any notable seasonal patterns in Tang Kristensen's collecting, we focused on monthly travel statistics. Unfortunately, the data we had available to populate the date field was insufficient for this task and we were forced to add a month field. Working from the memoirs, it turns out that the month of a certain stop can always be determined, even if the full date is missing. Once we created and populated this field, we were then able to calculate various monthly statistics (Figure 10).
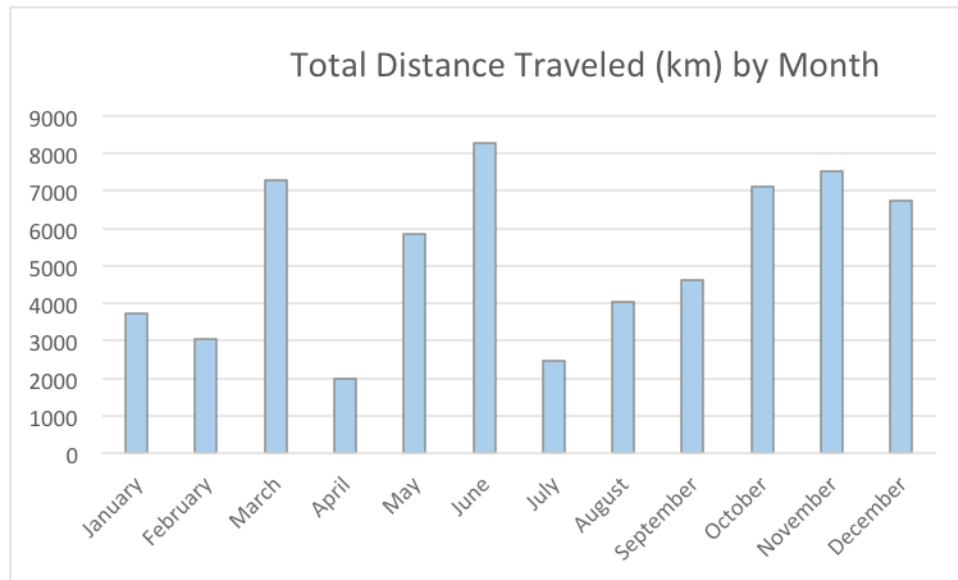
Figure 10: Distance traveled by month.

The most useful of these was the distance traveled by month. Not surprisingly, the month of June, with the longest days of the year, had the largest total aggregate distance. March, October and November were also represented as high travel distance months, with these months aligning with various school holidays. It seems somewhat surprising that December was so frequently represented, although it may be that these trips also align either with the Christmas holiday or capture the school holidays of late November to early December. Importantly, this relatively crude measure does suggest that the long dark nights of early winter were the most productive season for collecting, as the long days of summer could not compete with the indoor work and relative lack of activity that characterized the winter months on Danish farms. By mid to late winter, the weather was clearly a barrier to fieldwork. In future work, we intend to consider the seasons over longer periods of collecting, most likely based on Tang Kristensen's place of residence. Similarly, it may be worthwhile to look at *where* he traveled during these different seasons.

Area and population representation

One of the challenges of exploring fieldwork practices is the complex relationship between fieldwork sites and the expectations of the fieldworker. Over the years, Tang Kristensen began to favor certain places over others, revisiting these places multiple times. Curious as to where these repeated visits occurred, we decided to create a visualization of places where he made repeated visits (Figure 11). One can, in subsequent visualizations, correct for places that he lived, removing those from the map.
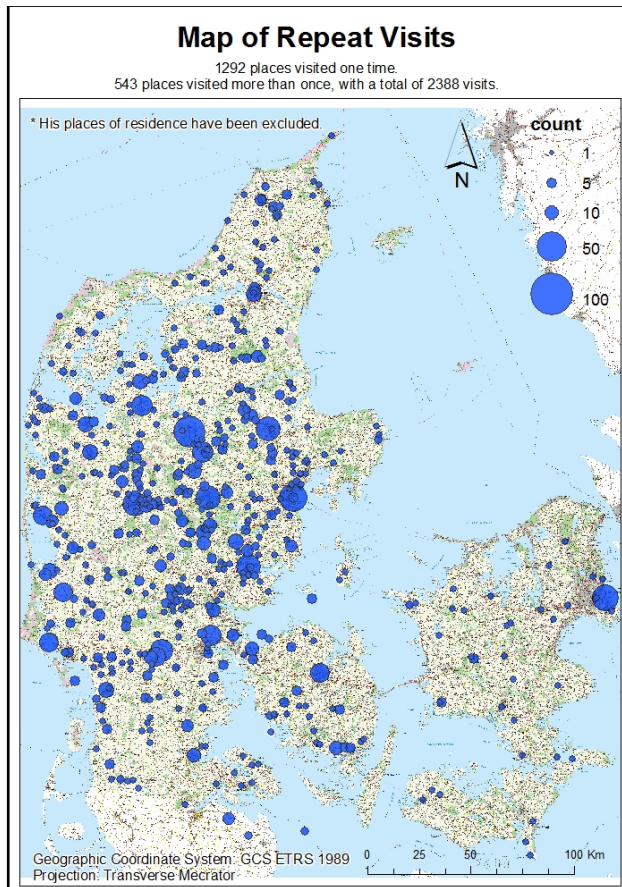
Rank	Place	Count
1	Viborg	58
2	Aarhus	46
3	Kjøbenhavn	40
4	Silkeborg	35
5	Randers	35
6	Herning	34
7	Horsens	32
8	Kolding	30
9	Brandstrup	27
10	Skive	26
11	Stovstrup	23
12	Odense	22
13	Ringkjøbing	21
14	Ulfborg	20
15	Vejen	19
16	Hurup	19
17	Ribe	15
18	Holstebro	15
19	Struer	15
20	Hammerum	15
21	Rødkjær	15
22	Give	14
23	Jebjærg	13
24	Kjellerup	13
25	Ry	13

Figure 11: Repeat visit count for the different places visited accompanied by the top places in a table.

Interestingly, the places he visited most frequently were cities or market towns, Viborg, Aarhus, Copenhagen, Silkeborg, Randers, Horsens, etc., the places where he either studied or had business or family interests. He did not collect to any appreciable degree from people in these places. One has to look down to the eleventh position in the rank list, discounting places where he resided, before one finds repeat visits to field collecting sites. This ranking of repeat visits treats individual spots separately and thus misses Tang Kristensen's intense focus on parishes (*sogn*) and small places within parishes.

In a second series of maps, we aggregate stop data at the level of parish and district (*herred*). By using data from the DigDag project, we are able to calculate the density of stops per parish and, given the areas of those parishes, the density of stops per square kilometer per parish. We then visualize these maps as choropleth maps, shading the areas according to either the total number of stops (Figure 12a) or the total number of stops per square kilometer (Figure 12b). These maps paint a somewhat different picture of where Tang Kristensen visited, as opposed to where he traveled, as shown in the repeat visits map.

Evald Tang Kristensen Collections 1868-1916
Number of Visits per Sogn
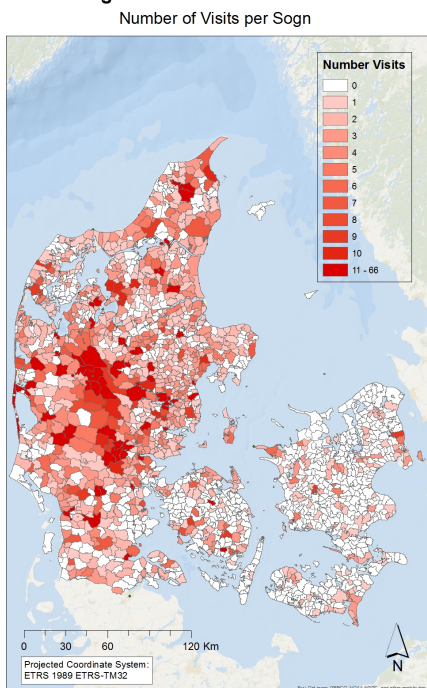
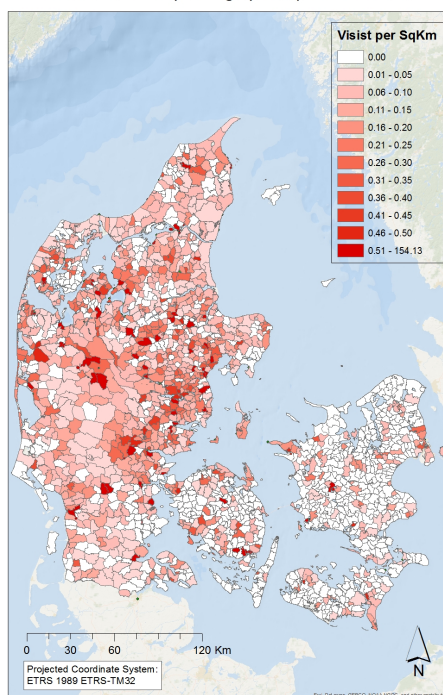Number of Visits per Sogn per Square Kilometer

Figure 12: Visits per parish (left), and visits per parish normalized by the area of the parish (right).

The unweighted choropleth reveals a significant number of repeat visits to parishes in the central part of Jutland, with a smaller grouping of equally visited parishes in the far northern provinces. Once corrected for parish size, the stop data becomes more even, with an emerging region in the south-east, an interesting corrective to Tang Kristensen's pronouncements that a great deal of his collecting was based in west Jutland.

There is a danger that distribution maps can overrepresent an area simply because it is densely populated: one need only think of Twitter maps that often map the location of cities, rather than the phenomenon they purport to map, to understand this problem. Of course, it would make little sense for a folklore collector to visit places where no one lived, and so these places are often rightfully underrepresented in folklore collections. In most representations of fieldwork activity, however, there is little attempt to capture the correlation between population density and productive fieldwork sites. To address this problem, we generate another set of choropleth maps based on parish (*sogn*) and district (*herred*) level population data (Figure 13).
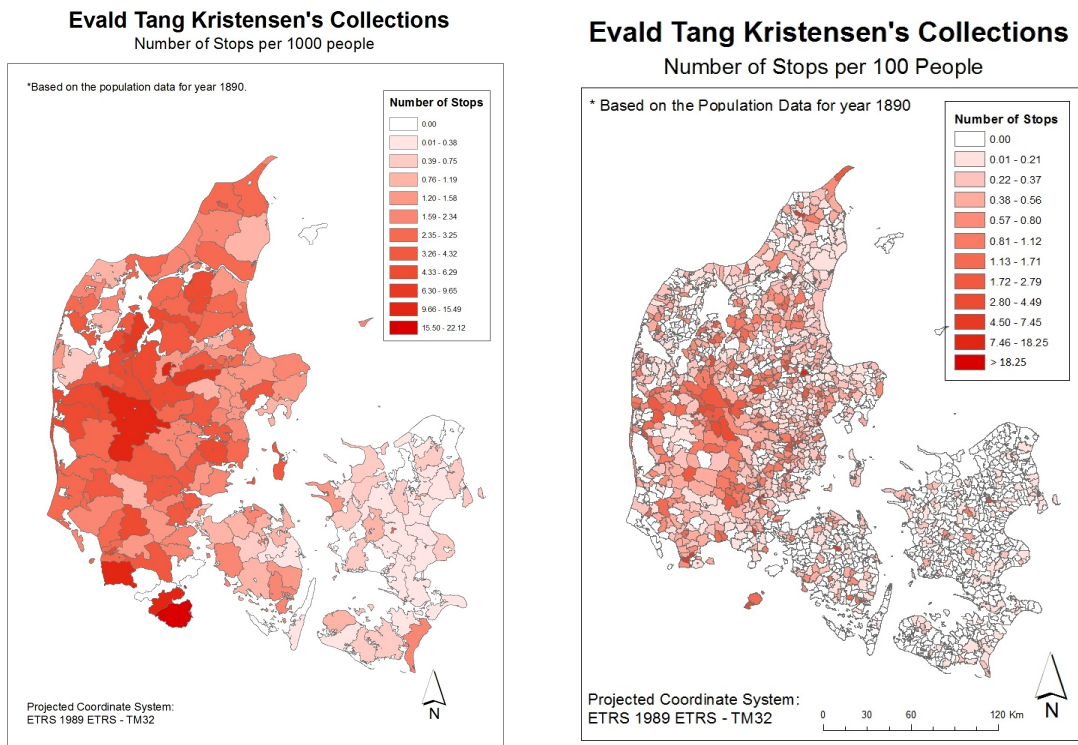
Figure 13: Stops per population by district (*herred*) and parish (*sogn*).

In these final choropleths, one begins to discern the underlying idea behind his evaluation of his collection as largely west Jutlandic: given the population density of west Jutland, he visited it far more frequently than its population would otherwise justify. That said, other parts of Jutland receive far more absolute representation in his collection.

Conclusion

This brief exploration of the fieldwork practice of Tang Kristensen reveals the benefits of using GIS systems to address complex historical questions, an approach known as historical GIS (hGIS). Although fieldwork practice has been addressed in recent years (Christiansen 2013; Österlund-Pötzsch 2013), that work is largely qualitative and descriptive. Here we provide fine-grained quantitative data to support some of the conclusions of folklore historians. We also provide consistent statistical measures to support the exploration of collecting practices and the collection itself. Importantly, these measures that takes into consideration the physicality and spatial dimensions of fieldwork and collecting.

Once we have route and route segment data, as well as stop and stop sequence data, and once we align this data with dates of collection and secondary data resources such as population data, there are myriad ways to address Tang Kristensen's fieldwork practice and,

by extension, the ongoing use of folklore as a means for imagining the nation. We believe this pilot project could be extended to other Nordic countries and collectors, and speaks to the need for ongoing collaboration not only across the region but also with other similar collections throughout Europe and beyond.

Our work reveals the shifting parameters of Tang Kristensen's field collecting, from his intensely local focus early on in his career to his more expansive and confident travels at the end of his career, when his collecting was no longer aligned with Romantic nationalist goals but was more in tune with a thick descriptive approach to Jutlandic rural life. By using hGIS techniques, we can provide a degree of detail about his travels that are missing from earlier studies. Our approach enables a truly macroscopic approach to folklore collecting, allowing us to interrogate Tang Kristensen's field collecting at varying levels of resolution. For example, we can move from the micro-consideration of a single fieldtrip, to a meso-consideration of all trips that included a particular parish, to a macro-consideration of all his trips taken as a whole.

REFERENCES

Anderson, B. (1983). *Imagined communities: Reflections on the growth and spread of nationalism.* New York and London: Verso.

Asbjørnsen, P.C., & Moe, J. (1842). *Norske folkeeventyr.* Oslo: Johan Dahl.

Bendix, R. (1997). *In Search of Authenticity: The Formation of Folklore Studies.* Madison: University of Wisconsin Press.

Christiansen, P.O. (2013). *Tang Kristensen og tidlig feltforskning i Danmark. National etnografi og folklore 1850-1920*. Copenhagen: The Royal Danish Academy of Sciences and Letters.

DigDag: *Digital Atlas of Denmark's Historical-Administrative Geography*. Last accessed February 18, 2017. Retrieved from http://www.digdag.dk

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik* 1: 269–271.

Dundes, A. (1977). Who are the folk? *Frontiers of Folklore* 5 (1977): 17-35.

Grimm, J., Grimm, W. & Grimm, F. (1816). *Deutsche Sagen*. Vol. 1. Berlin: Nicolaischen Buchhandel.

Grundtvig, S. *Danmarks gamle Folkeviser.* Copenhagen: n.p., 1853-1976.

Holbek, B. (1979). Folkemindevidenskab. In, *Københavns Universitet*. Edited by Jensen, P.J. Pp. 49-85. Copenhagen: University of Copenhagen.

Hult, M. (2003). *Framing a National Narrative: The Legend Collections of Peter Christen Asbjørnsen*. Detroit: Wayne State University Press.

Hyltén-Cavallius, G.O. (1944). *Svenska folksagor og äfventyr.* Stockholm: A. Bohlin.

Høje målebordsblade. (1842-1899).

Retrieved from http://download.kortforsyningen.dk/content/dtkhøje-målebordsblade

Kristensen, E.T. (1923-7). *Minder og Oplevelser.* Volumes 1-4. Viborg: Forfatterens forlag.
Krohn, K. 1926. *Die folkloristische Arbeitsmethode*. Oslo: P. Aschehoug.

Lönnrot, E. (1849). *Kalevala*. Helsinki: Suomalaisen Kirjallisuuden Seuran kirjapainossa.

Palmenfelt, U. (1993). *Per Arvid Säves möten med människor och sägner : folkloristiska aspekter på ett gotländskt arkivmaterial.* Stockholm: Carlssons.

Rockwell, J. (1981). *Evald Tang Kristensen: A Lifelong Adventure in Folklore.* Aalborg: Aalborg University Press.

Tangherlini, T. R. (2013a). *Danish Folktales, Legends and Other Stories*. Seattle: University of Washington Press.
Tangherlini, T. R. (2013b). The Folklore Macroscope. *Western Folklore* 72(1): 7-27.
Tangherlini, T. R., & Broadwell, P.M. (2014) Sites of (re) Collection: Creating the Danish folklore nexus. *Journal of Folklore Research* 51.2: 223-247.
Thiele, J.M. (1817). *Prøver af danske folkesagn.* Copenhagen: n.p.

Ward, D. (1981). *The German Legends of the Brothers Grimm*. Philadelphia: Institute for the Study of Human Issues.

Österlund-Pötzsch, S. (2014). Bodies in Motion: The Peregrinations of Early Finland-Swedish Folklore Collectors. *Journal of Folklore Research* 51(3): 253-276.