Digital Folkloristics: the Use of Computational Methods in Revealing the Characteristics of Folkloric Communication

Panel Topic

Folkloric communication is a specific mode of communication that is using existing knowledge and creatively adapting it in performance to a particular situation, audience, and intent. It always includes a portion of repetition and a portion of innovation or, in other words, as Michel de Certeau expresses it in his study of everyday practices, the main attribute of cultural transmission is the changing of everything that is being passed on (1984)\(^1\). The questions of authorship, stylistics and variation in time and space have been discussed and problematized in folklore studies for over a hundred years. The digital era has brought along the mass digitization of cultural heritage documents and the compilation of folklore databases and text corpora\(^2\). Yet, the use of computational methods in researching folklore and its specifics has thus far been more modest. The intention of this panel is to discuss the possibilities that computational methods have to offer to reveal the inherent qualities of folkloric communication represented in various text collections and research corpora. The participants of the panel deal with different source data including archival records and publications, as well as contemporary social media, and different aspects of folkloric communication and variation, contributing together to its general understanding.

The Panel: speakers and contributions

In this panel, the participants will introduce their projects, data, methods and research results to contextualize and elicit a discussion on the state-of-the-art in digital folkloristics, and on possible ways forward.

**Mari Sarv** and **Risto Järv** (Estonian Literary Museum) are going to analyse the essence of **folkloric variation** relying on the text corpora based on the collections of Estonian Folklore Archives. The database of Estonian Folksongs\(^3\) contains ca. 100,000 poetic texts, and the database of folk tales consists of 11,000 tales, both together with metadata. The studies of folksongs have shown that the statistical analysis of poetical features as well as the content parts of the allows us to locate the geographical origin (i.e. belonging of the songs to the tradition of a local community) quite precisely. At the same time there are clear differences in the geographical spread of linguistic-poetical features compared to the content elements. The stylistic analysis of folklore texts enables us to find out in how much the personal style of performer reveals in the recordings of traditional plots (‘types’ in the folkloristic discourse).

---


potential of computational methods to tackle the dichotomy of stability and variation in the folkloric communication intrigues the presenters most at the moment.  

Kati Kallio (Finnish Literary Society) will discuss the characteristics of oral poetry, including complex patterns of variation and tricky definitions of authorship, and concentrate on the challenges posed by the SKVR-corpus of Finnic oral poetry. The corpus represents various languages and dialects, orthographies, personal writing styles and traces of different modes of performance, but is held together by similar poetic registers. For this kind of specific poetic register in several related small languages and across a wide variety of genres, no ready-made tools for computational linguistic analysis exists. On the other hand, the corpus already includes a detailed thematic index, and the researchers have applied various manual methods to the corpus for hundred years. What kinds of new questions could be answered with computational methods? She will present some test analyses and discuss the future possibilities of digital folkloristics on oral poetry.

Greta Franzini and Emily Franzini (Göttingen Centre for Digital Humanities) will elaborate on two research projects focusing on the computational interrogation of folktale collections and corpora. Using the Brothers Grimm’s Kinder- und Hausmärchen as a case study and base reference, one project addresses the popularization of fairy tale motifs triggered by the Brothers, and seeks to algorithmically crawl web corpora to study the global network of motifs; the other examines the textual evolution of the Kinder- und Hausmärchen, starting with the first edition published in 1812 to the seventh and last in 1857. The number of fairy tales grew with every edition, and the numerous changes the Brothers made over the decades in terms of both style and content were symptomatic of societal interest and development. These seven editions represent an ideal testbed not only to computationally verify existing research about this progression but also to identify and distinguish the authorial and stylistic fingerprints of Jacob and Wilhelm Grimm. The discussion will demonstrate the possibilities afforded by the Digital Humanities to conduct web-scale and Big Data research.

Liisi Laineste (Estonian Literary Museum) is applying the methodologies of Digital Humanities to the folkloric aspect of social media content, pairing the results with theories of global information flow, participatory journalism and humour theory. Internet has become central in contemporary cultural communication – most of online communication can be treated as folklore, in which shared norms and values are constructed through cultural artifacts. Forums, commentary boards of news sites, blogs, Twitter and other social media applications have become commonplace during the last fifteen years. In the Internet, news, ideas and opinions travel fast. Social media spreads folklore in unforeseen volumes across national and cultural boundaries. Above all, the focus is on producing and consuming texts, images and multimedia. Social media is often perceived as a catalyst and accelerator of public discussion and citizen movements.

---