**MCDONALD — Crossroads**

<0 images>

**At the Crossroads of Data and Wonder: Algorithmic Approaches to Fairy Tales on Television**

McDonald, J. L. and Rudy, J. T.

Folktales and fairy tales are very often figured, in literary studies, as texts with sociohistorical and ideological force; for example, a given study might theorize how a particular discourse of fairy tales replicates gender roles, how fairy tales in American life support a Disnified culture industry, and so on. But significantly, fairy tales have also long been studied as data objects; since the 19th century, they have been conceptualized and categorized in terms of genre, origins and historic-geographic distribution, and form and structure—in essence, there is a strong sense that fairy tale scholarship depends on the cataloging of tales into various systems of ontological understanding.

In addition to existing in the written records and datatables of cataloguers, fairy tales are quite often associated with many other forms of transmission and dissemination. Certainly, it is common to encounter fairy tales (in many forms) in the context of oral storytelling, but tales also have been performed as plays and ballets, turned into cinematic or televised productions, illustrated in picture books and collections, and in our modern consumer culture, made into toys, dolls, video games, and other products. While the status of tales as texts has been long studied, these different medial iterations of tales have received less attention from scholars trained in literary studies and folklore programs. Yet it’s in this space, beyond the literary representation of fairy tales, where our modern comprehension of their power becomes most fully realized, existing and replicating in a way that Werner Wolf has described as ‘intermedial’. Intermediality, according to Wolf’s conceptualization, encourages attention to semiotic modes (speaking, writing, visually representing) and technological channels (printing presses and books; cameras, projectors, and films; broadcast stations, television sets, and TV shows) (2011). This notion of seeing narrative progress along various media channels is an instantiation of what Marie-Laure Ryan advocates. Ryan has suggested that the studying of narrative occurs best across media because narrative study that is focused only on language leaves out the nuances of different communicative modes and the technologies and social relationships that make such communication possible (2004). In other words, our premise is that studying intermedial forms of fairy tales help us better comprehend how these tales continue to work from telling to telling.

When it comes to folklore studies, the natural attraction to ontology and catalogue makes the digital humanities an attractive scholarly cohort, and a subfield often labeled ‘digital folklore’ is gaining traction. Because of the inherent ability of digital analysis methodologies to cross media (or deal with texts that exist in multiple media at once), it makes sense that, if we want to study intermedial fairy tales as a new entry point into these texts, then the crossroads of the computational (or ‘data-driven’) and the narrative is the space for fairy tale scholarship in the 21st century. Timothy R. Tangherlini has encouraged the use of algorithmic approaches to help better navigate the complex intermedial spaces of folklore narrative; leveraging the term ‘computational folkloristics’ (as distinct from digital folklore), Tangherlini states the importance of generating ‘algorithmic methods to assist in the interpretation of relationships or structures in the underlying data’ (2012). Indeed, Tangherlini’s efforts in helping give rise to the study of folklore through computational methods is the launching point for our project in looking at intermedial fairy tales;1 given the rich history of structural, ontological-based (or perhaps more appropriately termed, for our purposes, metadata-based) approaches to fairy tale study, the power of exploring fairy tales in intermedial spaces, and the concept of applying paradigmatic methodologies from the digital humanities as an analytical tool for analyzing fairy tales today, we have begun to identify relationships between tales, media, history, geography, production and reception, and scholarly and popular audiences in a new light.

For the bounds of this proposal (which is one piece of a recently awarded SSHRC collaborative grant to look at intermedial fairy tales), we are concerned with the occurrence of fairy tale storytelling in a very specific intermedial space—that of modern American television. For the past 85 years, fairy tales have found their way into television broadcasts in numerous ways, whether it be quick allusions to fairy tales to advance otherwise non–fairy tale narratives; spectacular retellings of tales as live events or made-for-TV ‘specials’; fairy tale contexts for particular episodes of traditional, procedural television; longer-running arcs of serialized shows; or even premises for entire shows constructed on fairy tale tropes. We have begun with a teleography—a metadata-rich bibliography of television fairy tales compiled by Kendra Magnus-Johnston and recently published in *Channeling Wonder: Fairy Tales on Television*. With over 1,000 data points (individual episodes or TV specials related to a fairy tale in some way), our project will approach this metadata in ways similar to how historical fairy tale scholars have: looking for patterns in terms of narrative similarity, common genre, chronological proximity, and so forth, but doing so with tools and experiments that leverage data processing and visualization methods that have become significant paradigms in digital humanities scholarship.

Our goal in beginning with data modeling is to reposition the existing teleography as a sort of intermedial fairy tale corpus that can be mined and analyzed visually, spatially, and temporally. Through automated text parsing, processing, and categorization; aggregating information both from the teleography as well as from other sources (semantic repositories of linked data such as Freebase.org and other relational datasources such as the crowdsourced website and API thetvdb.org); and generation of our own dynamic dataset (serialized in various forms—tabular CSV and JSON right now), the scholars and students involved in this phase of the project have been better positioned to comprehend the scope of the fairy tale data and metadata, and identify the types of interactions and relationships that might be exposed by different types of visualizations. In many ways, this sort of data structuring is a reimagining of the work done by Antti Aarne and Stith Thompson (of the Aarne-Thompson classification system) or by Hans-Jorg Uther’s *ATU Catalogue*; in fact, our data model, mastered in a MongoDB instance and published out to a web app, uses a fairy tale’s ATU number as a unique identifier, supplementing missing tales as needed (the current state of the searchable, filterable teleography web app is available at http://fttv.byu.edu/#datatable). This preliminary data work has already demonstrated ways that the medium of television has gone far beyond the traditional animal, magic, and folktales of the ATU ontology; we have found it useful to incorporate into our data such literary “fairy tales” as *Peter Pan* or *The Wizard of Oz,* Hans Christian Andersen tales such as ‘The Ugly Duckling’, ‘The Little Mermaid’, and ‘The Snow Queen’, or more abstract references such as ‘princess-ness’, ‘grimm-ness’, and so forth. Our plans for the underlying data source are lofty, as we are continually working to incorporate more and more occurrences of fairy tales on television as well as to augment the data with more metadata, such as broadcast dates, actors involved, broadcast and cable channels where they occur, and so forth.

Of course, given that our goal is to approach the study of fairy tale narratives in an intermedial space, we aren’t satisfied with just developing our own indexed data source; rather, we have begun to analyze its structured form by generating a collection of visual representations of the data (see http://fttv.byu.edu/#visualizations). With everything from simple histograms, charted occurrences of fairy tales on TV over time, and force-directed or matrix-based network graphs built with d3.js (that can juxtapose different metadata variables with our expanded ATU tale-type ontology), we are recognizing new patterns and insights about the tales themselves as well as about the medium in which they’ve appeared. For example, one student, seeing the highest frequency of fairy tale appearances in American television to be ATU 510A (Cinderella), is now investigating the relationship between the narrative of the servant-turned-princess and the national discourse of American exceptionalism. Another is looking at the rise of reality television as a replacement for the traditional fairy tale story, while a third is exploring the many references in a single television series (*Fractured Fairy Tales*, snippets of tale retellings which appeared in 91 different episodes of *The Rocky and Bullwinkle Show*) as an entry point for understanding commodity capitalism and self-valuation during the late 1950s–early 1960s.

Cristina Bacchilega has suggested that modern treatment of fairy tales expresses the evolution of human hopes and fears (1997); television’s ability to appropriate, repurpose, or refigure tales confirms this. Combining multiple tales to tell new stories is also leading scholars to look at intertextuality between fairy tales.

**Note**

1. As far as we can tell, there have not yet been any computational approaches to intermedial fairy tale analysis. There are a few projects in which scholars have sought to apply digital humanities methodologies to traditional fairy tale texts such as *Grimm’s Fairy Tales*; notable among these are Declerck et al., who describe a TEI-based markup format and a tool for semiautomated markup of fairy tale text (2011), and Weingart and Jorgensen, who use algorithmic analysis of European Fairy Tales to discuss conceptualizations of the body (2013).

**References**

**Bacchilega, C.** (1997). *Postmodern Fairy Tales: Gender and Narrative Strategies*. University of Pennsylvania Press, Philadelphia.

**Declerck, T., Scheidel, A. and Lendvai, P.** (2011). Proppian Content Descriptors in an Integrated Annotation Schema for Fairy Tales. *Language Technology for Cultural Heritage: Selected Papers from the LaTeCH Workshop Series*, Springer, pp. 155–70.

**Greenhill, P. and Rudy, J. T.** (eds.). (2014). *Channeling Wonder: Fairy Tales on Television*. Wayne State University Press, Detroit, MI.

**Ryan, M.-L.** (2004). *Narrative across Media: The Languages of Storytelling*. University of Nebraska Press, Lincoln.

**Tangherlini, T., Abello, J. and Broadwell, P.** Computational Folkloristics. *Communications of the ACM,* **55**(7) (2012): 60–70.

**Weingart, S. and Jorgensen, J.** Computational Analysis of the Body in European Fairy Tales. *Literary and Linguistic Computing,* **28**(3) (2013): 404–16.

**Wolf, W.** (2011). (Inter)mediality and the Study of Literature. *CLCWeb: Comparative Literature and Culture,* **13**(3) (2011), <http://dx.doi.org/10.7771/1481-4374.1789>.