# Mining Motifs in Fairy Tales

**Saba Shahsavari, Alessa Mayer, and Fabian M. Suchanek**
Telecom Paris, Institut Polytechnique de Paris, France

## Abstract

A motif is a pattern that appears across stories, tales, or legends, such as "a prince falls in love with commoner". We propose a method based on large language models to extract such motifs automatically from stories. Our approach addresses challenges such as motif synonymy, and achieves high alignment with expert-annotated ground truth. We illustrate the utility of our extraction in downstream applications such as mitigating gender bias in existing motif collections, studying narrative event sequences, quantifying cross-regional story similarity, validating the correlation between linguistically related cultures, and tracing the relatedness of religious traditions.

## 1 Introduction

Stories, from fairy tales to myths and religious legends, encode cultural values and narrative structures that persist across regions and eras. Some of these stories share basic ideas even if they are told in different societies. For example, the idea of a prince marrying a poor girl despite social barriers appears not just in the European Cinderella story, but also in the Malay-Indonesian Bawang Putih Bawang Merah tale and the Chinese Ye Xian story. Such plot elements (or *motifs*, as they are known (Thompson, 1955–1958; Acharya et al., 2024)) can tell us a wealth about the host culture of the story and its link to other stories. For example, it is known that two linguistically related cultures are more likely to share motifs, even if they are geographically distant (Da Silva and Tehrani, 2016). One instance of this phenomenon are religious mythologies, which share concepts when the religions are descendants of each other or of a common ancestor (Noyer, 2011; Mallory and Adams, 2006; Suchanek, 2026). Beyond that, plot motifs can tell us about prevalent preoccupations of people and the differences and similarities of these preoccupations across cultures. If two cultures are found to share such shared cultural heritage, this could potentially even help them to view each other more favorably.

Indeed, several studies have analyzed such motifs in stories across cultures. Particularly large works are the groundbreaking studies by Aarne (1910), Thompson (1955–1958), and Uther (2011), as well as Berezkin and Duvakin (2016b). However, all of these studies rely on a manual identification of motifs. Such an analysis is not just labor-intensive, but it also limits any analyses of motifs to corpora where such annotations happen to have been made.

In this paper, we propose to extract such motives automatically from stories. This is a challenging endeavor, because we must address not just the variety of tale types, but also geographic context, and motif synonymy (e.g., between "Dream prompts a journey" and "Dream makes protagonist travel"). This is why, to the best of our knowledge, no large-scale automated extraction of motifs has yet been undertaken. Our idea is to capitalize on the recent advances of large language models (LLMs). We design prompts to extract motifs, and clustering techniques to group them together (Figure 1). Our experiments show that our methods can extract, with high similarity, the motifs that have been manually determined by experts. Furthermore, our motifs predict the story type better than the text alone, indicating that our motifs capture the essence of the story without being distracted by its elaboration.

We then show how our extracted motifs can be used to mitigate gender bias in existing motif collections, visualize common sequences of story events, determine the similarity between stories of different geographical regions, validate the finding that cultures share more motifs if they are linguistically related even if they are geographically distant (Da Silva and Tehrani, 2016), and show relatedness between religious traditions.

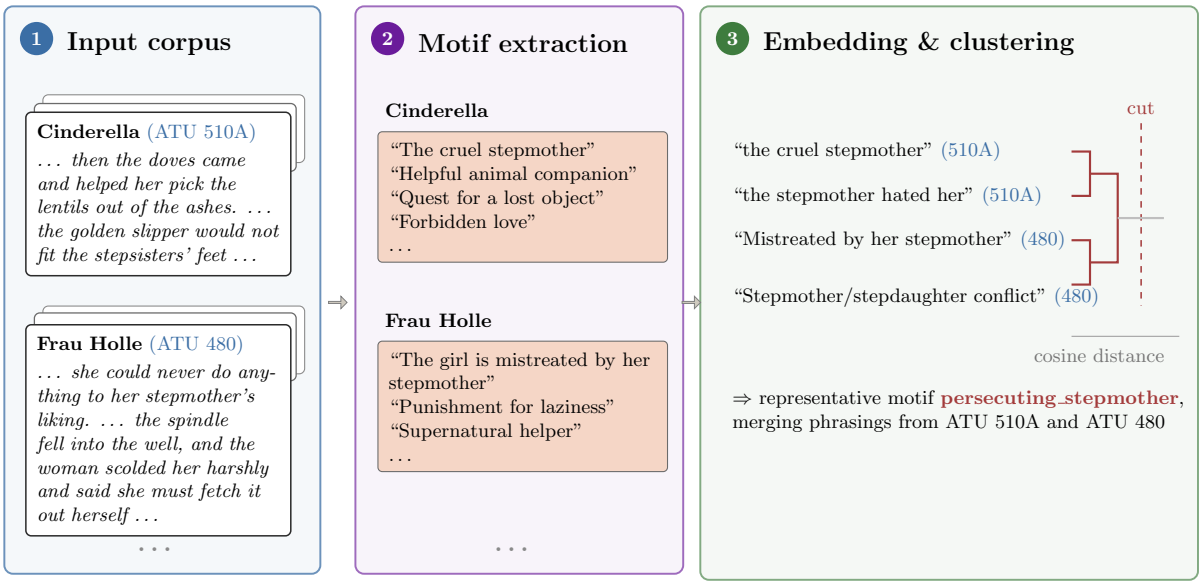


Figure 1: Overview of our approach. Motifs are extracted from folktales by a large language model, encoded with a pretrained sentence embedder, and clustered into a normalised motif inventory that recurs across tales and ATU types. These motifs form the basis of all analyses in this paper.

## 2 Related Work

### 2.1 Evolutionary and Psychological Perspectives on Narratives

Stories are not merely cultural artifacts, but fundamental to the human experience. Jonathan Gottschall (2012) describes humans as "storytelling animals", suggesting that narratives act as a central cohering force in human life. Stories can establish shared identity within groups, transmit ancestral knowledge and cultural norms, serve as sources of metaphor, and provide role models for navigating difficult situations. Beyond these functions, stories offer narratives of archetypal human situations, functioning as a "flight simulator" for the mind and enabling individuals to rehearse responses to negative emotions and social challenges in a safe environment. Furthermore, stories facilitate perspective-taking (Pinker, 2012), a phenomenon Suzanne Keen (2013) terms *narrative empathy*. This capacity allows readers to broaden their horizons, encounter unfamiliar viewpoints, and develop empathy toward people whose experiences differ from their own.

### 2.2 Manual analysis of motifs

**Collection of stories.** Early collectors and authors have contributed significantly to the enduring canon of fairy tales and their characteristic motifs, in particular Giambattista Basile (1634–1636) with *Lo cunto de li cunti*, Charles Perrault (1697) with *Contes de ma mère l'Oye*, Jeanne-Marie Leprince de Beaumont (1798) with *Magasin des enfants*, and Grimm and Grimm (1812) with their fairy tale collection.

**Classification of stories.** Early efforts in fairy tale classification were pioneered by Antti Aarne (1910) in his *Verzeichnis der Märchentypen* and subsequently expanded by Stith Thompson (1961) in *The Types of Folktale*. Building upon this foundational work, Hans-Jörg Uther (2011)'s revised and expanded *Types of International Folktales* further refines the classification system, giving rise to the Aarne-Thompson-Uther (ATU) index. It organizes folktales into seven broad categories, namely Animal Tales, Tales of Magic, Religious Tales, Realistic Tales, Tales of the Stupid Ogre, Anecdotes and Jokes, and Formula Tales. Within each category, folktale types are further subdivided by motif patterns until individual types are listed, resulting in a total of around 2400 ATU types, such as ATU 510A (Cinderella), ATU 333 (Little Red Riding Hood), and ATU 300 (The Dragon-Slayer).

The structuralist approach of Vladimir Propp (1968) in his *Morphology of the Folktale* is complementary, as it focuses on the sequence of functions within a narrative.

**Collection of motifs.** A *motif* is a distinctive, recurring, widely used idiom, often originating in folklore and usually strongly anchored to a particular cultural or national group (Acharya et al.,

2024). **Motifs thus go beyond ATU types by describing more fine-grained story elements.** Examples include "the helpful animal", "the wicked stepmother", and "the youngest sibling succeeds". The latter, for example, appears in both ATU 402 (*The Three Feathers*), where the youngest son alone succeeds at a series of tasks with the aid of an enchanted animal, and ATU 550 (*The Golden Bird*), where the youngest of three sons succeeds at a quest after his elder brothers fail, illustrating how the same motif can recur across distinct ATU types. Contemporary scholarship, as surveyed in works like Katrin Pöge-Alder (2007)'s *Märchenforschung – Theorien, Methoden, Interpretationen*, continues to recognize the importance of motifs in understanding both the common threads and the unique adaptations of fairy tales.

One collection of such motifs is Yuri Berezkin's *Analytical Catalogue of World Mythology and Folklore* (Thuillard et al., 2018), based on fifty thousand abstracts of texts from all over the world. However, this dataset maps each motif to its region, not to its story, and is thus less useful for our purposes. Thompson (1955–1958)'s *Motif-Index of Folk-Literature*, too, is a comprehensive catalog of motifs, ranging from magical objects and supernatural adversaries to specific tasks and character archetypes. Each of these motifs is mapped to an ATU type, making it a perfect source for our work.

### 2.3 Automated analysis of motifs

While the computational analysis of folk narratives has gained attention in recent years, direct approaches to narrative motif extraction remain underexplored. Existing research focuses primarily on higher-level classification tasks such as story type or genre identification, which, although related, do not entail the extraction of motifs as standalone narrative elements. For instance, Nguyen et al. (2013) address the automatic classification of folktales into story types according to schemas such as the ATU. Their method employs a learning-to-rank framework using a combination of information retrieval techniques, lexical similarity measures (e.g., TF-IDF, Jaccard), and structural features such as subject-verb-object (SVO) triplets. Although motifs are acknowledged as part of what defines a story type, the system does not attempt to extract them individually.

Similarly, Nguyen et al. (2012) focus on the automatic classification of folk narrative genres, such as fairy tales, urban legends, and jokes. The authors mention the use of lexical (e.g., character n-grams), structural (e.g., sentence length, punctuation), and metadata features (e.g., story keywords and summaries) to classify the stories. Though the approach is effective for genre prediction, it does not model or extract motifs as discrete entities.

Flaccavento et al. (2025) propose a method for detecting tropes in short social media posts using supervised learning. Unlike their classification-based approach, we identify motifs from scratch. We also analyze them semantically to uncover narrative structure and inter-story relationships.

Pucci et al. (2025) investigate whether LLMs process animacy similarly to humans. Like us, they aim to capture subtle semantic and narrative properties that go beyond superficial lexical associations.

## 3 Our Datasets

**AFT Dataset.** Among the publicly available ATU-annotated folktale collections, the most comprehensive and extensively annotated is the *Folktexts Collection* (Ashliman, 2021), maintained by D.L. Ashliman since 1996. What makes this collection particularly valuable for computational analysis is its consistent annotation with ATU tale type identifiers, which allows for large-scale classification and pattern analysis. We used the preprocessed version of the dataset provided by Hagedorn and Darányi (2022), who carefully extracted and annotated the tales from Ashliman's website and made the dataset publicly available. This dataset, the Annotated Folk Tales (AFT), comprises 1518 tales, which vary significantly in length, ranging from as short as 10 tokens to as long as 12406 tokens, with an average length of 979 tokens per story. The dataset maps each story to one of 182 ATU types, and so each ATU type has between 1 and 31 stories (8.3 on average). To ensure sufficient data for meaningful pattern discovery, we excluded ATU types with fewer than five stories. After this filtering step, the final dataset used for analysis consists of 1331 stories across 133 distinct ATU types. We augmented these ATU types with the associated motifs from the Thompson (1955–1958) motif index, with 332 mappings in total, i.e., each ATU type has 2.5 motifs on average.

**Yashpeh Dataset.** The Yashpeh (2013) dataset is a large-scale collection of folktales sourced from public-domain books and compiled from a wide range of cultural and geographical traditions. Unlike AFT, this dataset does not include ATU tale

type annotations. However, it provides rich metadata on narrative tradition, which enables coarse-grained cultural and regional analysis. While Ashliman's Folktexts Collection primarily focuses on Indo-European traditions, Yashpeh extends to a broader global distribution of folktales, incorporating narratives from diverse regions beyond the Indo-European sphere. This makes it particularly valuable for studying cross-cultural narrative structure without being constrained by pre-existing typological labels such as ATU classifications. We crawled more than 2000 folktales from the Yashpeh database and, after removing stories that overlap with the AFT dataset, retained 1852 unique folktales, each annotated with a list of one or more locations.

**Our story dataset.** The combination of both datasets yields 3183 stories. Of these, 1331 stories have ATU tale type annotations, covering 133 distinct ATU types, as well as motif annotations, with 2.5 motifs per story. Each story is further annotated with its geographical region.

**Our religious legends dataset.** To apply our method to a new scenario, we constructed a dataset of religious legends. For this purpose, we crawled all Wikipedia pages of the category *Religious Folklore* up to a depth of 3, obtaining 764 pages. We removed pages about movies, modern fiction, songs, and television shows by help of keywords, leaving us with 413 pages. Among these, we could map 372 pages to major religions by help of keywords. These pages contain 417 stories, which distribute as follows (multiple mappings possible): Hinduism (113), Christianity (98), Ancient Roman (58), Shintoism (56), Buddhism (40), Judaism (37), Islam (30), Chinese Folk (28), Ancient Greek (26), Taoism (21), and Jainism (11).[1]

## 4 Our Annotation

**Motif extraction.** We use a Llama 3 model (*llama-3.1-8b-instant*) with an all-mpnet-base-v2 embedder. We used the following prompt to extract motifs from stories:

*You are a computational folklorist analyzing narrative patterns in folktales. Extract the key recurring narrative motifs from the following story. A motif is a recurring, distinctive narrative element or pattern in a story. It can be a specific action, object, character type, supernatural event, or situational feature that appears across multiple tales and carries recognizable symbolic or functional meaning (e.g., "a forbidden door", "transformation by eating a magical fruit", "helpful animal companion", "quest for a lost object"). Return a JSON object with the following structure: {{ "Motifs": ["motif1", "motif2", "..." ] }}*

The prompt is followed by the story.

**Motif clustering.** After extracting motifs, we perform Truncated SVD dimensionality reduction on their embeddings, which projects them onto a lower-dimensional subspace that retains most of their variance. This follows the established use of SVD-based reduction in text analysis (Deerwester et al., 1990), and mitigates the degradation of cosine-based distances in high-dimensional spaces prior to clustering; prior work has further shown that such reduction can be applied to pre-trained sentence embeddings with little or no loss in downstream performance (Zhang et al., 2024). We then followed up by a hierarchical (agglomerative) clustering. This step serves two purposes. First, it accounts for the surface variation that is common in LLM outputs (the same motif is often phrased differently across extraction runs, and clustering merges these near-duplicate surface forms). Second, hierarchical clustering allows us to represent motifs at different levels of granularity: by varying the distance threshold, we can obtain coarser or finer-grained motif groupings, depending on the analysis. This threshold is defined in terms of the cosine distance between motif embeddings; we use a default value of 0.3, which merges near-duplicate motifs differing only in surface form, while higher thresholds yield progressively coarser motif groupings. At this default threshold, clustering yields 11808 fine-grained motifs on our story dataset, with an average of 9.2 motifs per story.

The *cluster centroid* is given by the mean embedding of all motifs in the cluster. To keep motifs interpretable, we select the most frequent motif in each cluster as its *representative motif*. This allows us to conduct downstream analysis and exploration at multiple levels of granularity, using either the representative motif (for interpretability) or the centroid (for numerical analyses). Figure 1 illustrates this process.

**Validation through ATU type classification.** To validate our method, we extracted the motifs on our

[1] Unfortunately, we were unable to harvest stories from the Bahai Faith, Zoroastrian, Sikhism, Indigenous Religions, and other extinct or existing traditions.

dataset, and used them to predict the ATU type of the tales. For this purpose, we represent each tale as a vector with one dimension for each motif, which we set to one if the motif is present, and to zero else. To ensure reliable evaluation, we restricted the experiment to ATU types with at least 15 stories. The resulting dataset was divided into training and test sets using a stratified 80/20 split, preserving the distribution of ATU categories across both subsets. The vectors were L2-normalized and used to train a multinomial logistic regression classifier with balanced class weights. Using only the motif-cluster histograms as input, the classifier achieved an accuracy of 79.5% on the held-out test set, showing that our motifs contain enough information to reconstruct ATU types. For comparison, a random baseline evaluated over 10 stratified samples with different random seeds reaches only 4.8% accuracy, while a logistic regression classifier trained on TF-IDF features of the raw story texts achieves 71.0%. This shows not only that our motifs retain all information needed to determine the type of story, but also that they capture only the essence of the story and not its elaboration, leading to a better classification result than the text itself.

**Validation against Thompson's Motif Index.** To assess the completeness of our extracted motifs, we performed a cross-reference analysis between our automatically extracted motifs and the corresponding gold-standard Thompson motifs. Since the two resources differ in wording and granularity, we compared motif embeddings rather than relying on exact lexical matches. Each Thompson motif was considered covered if at least one extracted motif had a cosine similarity of 0.70 or higher with its embedding. Our analysis showed that the extracted motifs achieved 90% coverage of the Thompson motifs associated with the corresponding ATU types. This high coverage indicates that the proposed extraction pipeline successfully recovers the vast majority of the canonical narrative motifs identified by human experts.

A cursory analysis shows that most of the remaining unmatched Thompson motifs are attributable to dataset incompleteness rather than failures of the extraction method. Specifically, the Thompson Motif Index is based on a broader set of tale variants than those available in our corpus, meaning that some motifs are associated with story variants that are absent from the AFT dataset used in our experiments. Consequently, these motifs could not be recovered by any extraction method operating solely on the available stories.

**Motif-based story retrieval.** To validate that our extracted motifs capture meaningful narrative structure beyond raw story text, we conduct an ATU tale type retrieval experiment. Given a query story, the task is to rank all other stories by similarity and evaluate whether stories of the same ATU type are retrieved near the top of the ranking, measured by Mean Average Precision (MAP), Precision@1, Precision@5, and Precision@10. We compare two retrieval systems. The baseline encodes each story as a dense vector using the all-mpnet-base-v2 sentence transformer, with overlapping chunk averaging for stories exceeding the 384-token limit. Our motif-based system represents each story as the set of embedding vectors of its extracted motifs and computes inter-story similarity via a symmetric max-pool strategy: For each motif in the query, we find its most similar motif in the candidate story and average these maximum similarities across both directions. This avoids collapsing motif identity into a single mean vector. Both systems are evaluated on the ATU-annotated stories across 24 tale types with at least 15 stories each.

Table 1 shows that our motif-based retrieval achieves MAP 0.447 versus 0.376 for raw text. i.e., a 19% relative improvement, with consistent gains across all metrics. These results show that automatically extracted motifs capture structural narrative properties that generalize across different surface realizations of the same tale type, beyond what is recoverable from raw text embeddings alone.

## 5 Using Our Extracted Motifs

**Mitigating gender bias.** Thompson's original index suffers from a bias towards male entities (Singh, 2023). To quantify this bias, we counted how often male and female indicators appear in the motifs (using "she", "her", "hers", "woman", "women", "girl", "girls", "mother", "daughter", "sister", "queen", "princess", "witch", "maid", "wife", and "widow" to identify women, and the male analogs to identify men). In the Thompson motif index, 17,8% of the motifs (weighted by number of appearances in the stories) are male-centered, and only half as many, 7,9%, contain female roles. In our newly extracted motifs, in contrast, gendered motifs split almost equally between male-centric (7,5%) and female-centric (7,1%). Our index thus corrects a gender bias that

| Method | MAP | P@1 | P@5 | P@10 |
|---|---|---|---|---|
| Raw text (sentence embedding) | 0.376 | 0.730 | 0.595 | 0.486 |
| Motif max-pool (ours) | **0.447** | **0.743** | **0.606** | **0.529** |

Table 1: ATU tale type retrieval results. 466 stories across 24 tale types (minimum 15 stories per type).

was present in Thompson's motifs, but possibly not in the stories themselves.

**Motif transitions.** To study the narrative logic connecting motifs within folktales, we constructed a directed transition graph over the extracted motifs. For every ordered pair of motifs $(A, B)$ we counted how often $B$ immediately followed $A$ within the same story, and normalized these counts into a conditional transition probability $P(B \mid A)$. Figure 2 shows the twenty transitions with the highest conditional probability. We find familiar patterns: a magical animal appears and helps in a quest for a lost object; a dream allows the truth to be revealed; and cleverness is used for trickery. Interestingly, many sequences end in the theme of redemption and forgiveness.

**Regional story-similarity.** To examine how motifs vary geographically, we computed the mean embedding of all motifs associated with each region, yielding region-level vectors that summarize the dominant narrative elements of each area, and projected them into two dimensions using t-SNE. For this analysis we discarded stories lacking clear region metadata as well as regions with fewer than ten stories, while deliberately retaining two collections whose origins remain debated, *Aesop's Fables* and *One Thousand and One Nights*, in order to test whether our analysis can offer new insight into their provenance.

Figure 3 shows that geographically and culturally adjacent regions are grouped together, indicating that motif usage carries a regional signal. Aesop's Fables offer a particularly instructive case: The fables are traditionally attributed to Aesop, a storyteller believed to have lived in ancient Greece around 620-564 BCE. They were latinized by Gaius Julius Phaedrus in the 1st century CE, and are generally considered European folklore. However, the embedding representing Aesop's Fables is positioned much closer to the Indian cluster than to other European regions. And indeed, several Aesopic fables closely parallel stories found in the Indian *Jataka Tales* and the *Panchatantra*. Scholars have noted over a dozen such narrative correspondences, often with thematic or structural alignment but variations in surface details (Berezkin and Duvakin, 2016a). Our motif-based analysis reflects this connection.

**Frequent motifs per tradition.** Figure 4 shows a phylogenetic tree of languages with their common story motifs. Our analysis suggests that "supernatural or magical events", "redemption and forgiveness", and "forbidden love" are universal motifs across all regions we studied. Previous research (Da Silva and Tehrani, 2016) has shown that linguistic relatedness between two cultures is a better indicator of the similiarity of their stories that geographic proximity. For example, we expect Romanian stories to be closer to Spanish stories than to Hungarian stories (even if Romania borders Hungary), because Spanish and Romanian share Latin roots. To confirm this hypothesis, we performed a Mantel test (Mantel, 1967) comparing pairwise linguistic distances with pairwise motif distances across provenances. The dataset was filtered to retain only unambiguous geographic provenances. This excluded, for instance, author attributions (e.g., Aesop), multi-origin collections (e.g., *1001 Nights*), and vague labels (e.g., Europe). Moreover, we filtered regions that included less than three tales to ensure stable motif profiles. This left 45 provenances comprising 885 tales. Linguistic distance was computed as a hierarchical tree-based metric, ranging from 0 (same language) to 4 (unrelated language families), derived from standard genealogical language classifications. Motif distance was measured in two ways: (i) cosine distance over TF-IDF vectors of the extracted motif descriptions per provenance, and (ii) Jaccard distance over the sets of unique motif strings. The Mantel tests were calculated using the mantel function from the scikit-bio library (*skbio.stats.distance.mantel*). Both Mantel tests (Spearman correlation, 9,999 permutations) yielded significant positive correlations between linguistic and motif distance (TF-IDF: r = 0.289, p = 0.018; Jaccard: r = 0.214, p = 0.017), confirming that provenances with more closely related languages share more similar motifs.

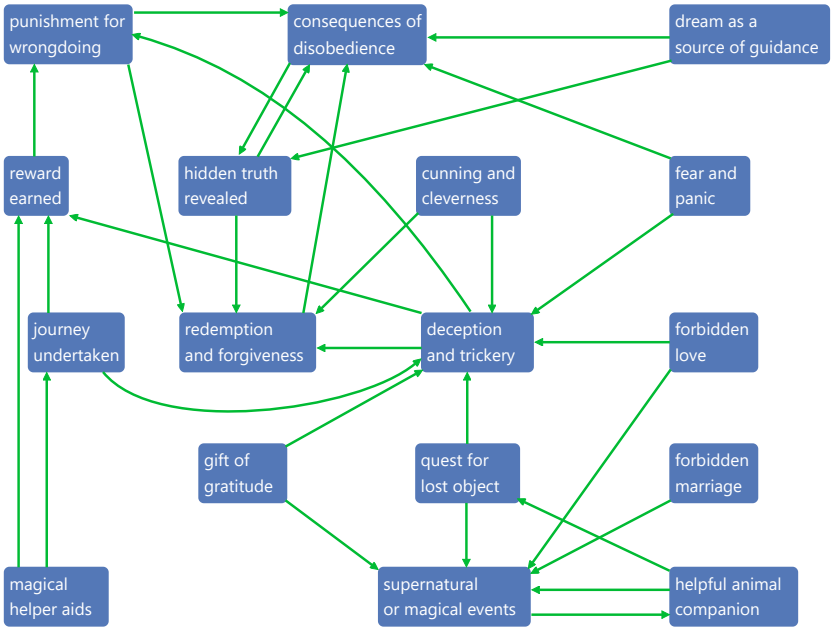


Figure 2: Motifs that frequently follow each other, showing the top 20 highest probability transitions.

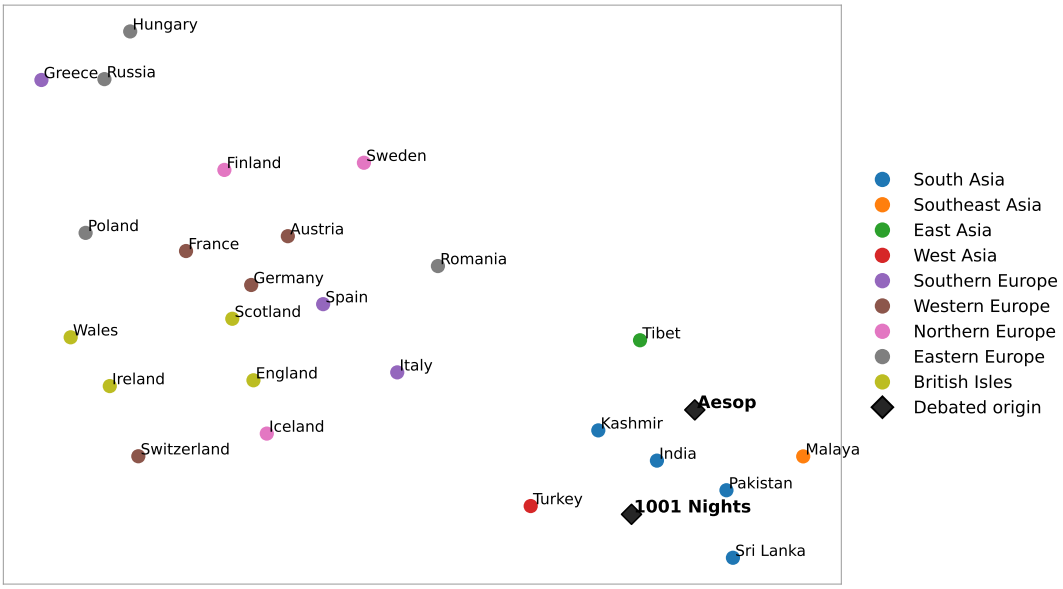


Figure 3: Regional narrative patterns. Each point represents a geographical region, based on the mean embedding of motifs from stories attributed to that region.

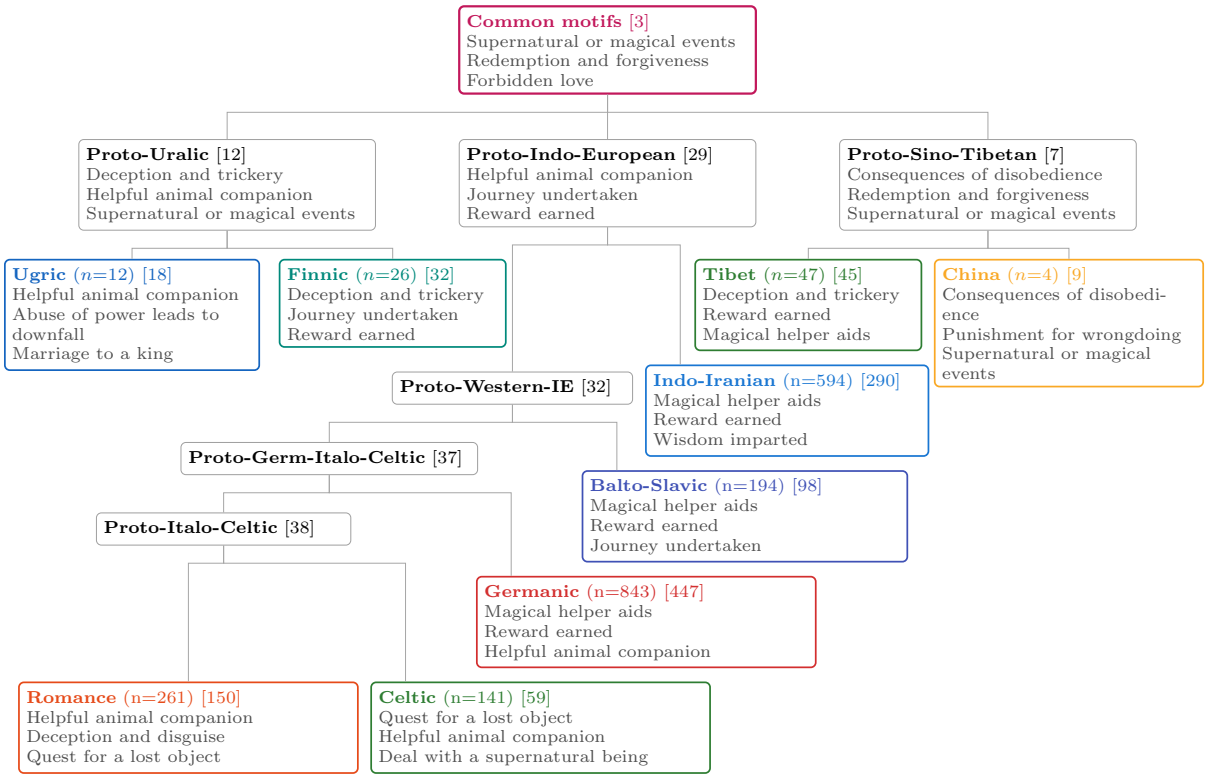


Figure 4: A phylogenetic tree of languages with their common story motifs. Each of the 5 leaf branches includes only motifs that appear at least 5 times among that branch's tales. Each ancestral node's motif set is the exact intersection of its children's sets. The 3 example motifs listed under each box are chosen randomly from the most frequent motifs. $n$ denotes the number of stories and the number in brackets denotes the number of motifs.

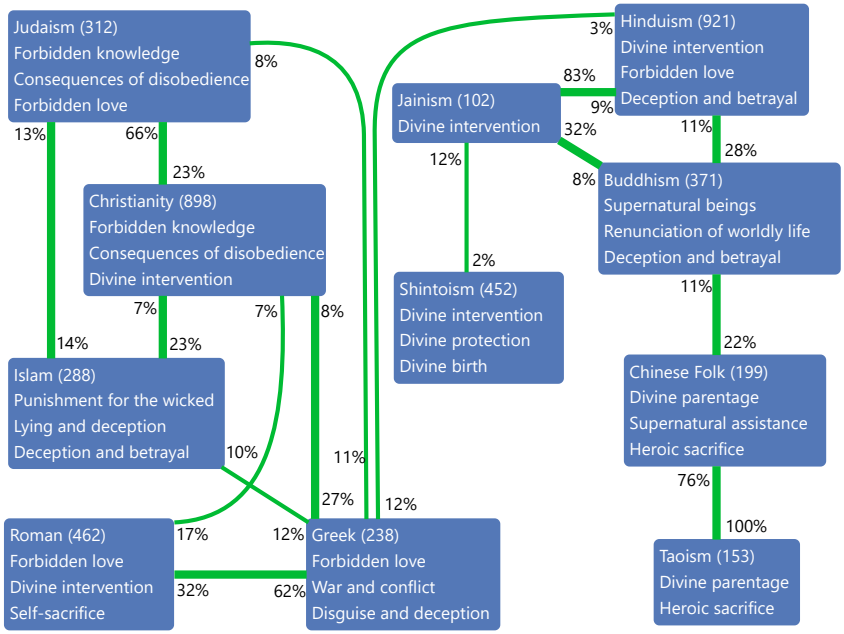


Figure 5: Number of motifs, most frequent motifs, and percentage of motifs shared with another story set (if at least one value is $> 10\%$). All numbers weighted by the number of occurrences of the motif in the story set. All numbers exclude the "universal" motifs.

To evaluate whether our motifs provide a superior signal compared to surface-level language, we conducted a control experiment applying the identical Mantel tests directly to the raw story texts. While raw text TF-IDF distance yielded a comparable correlation ($r = 0.289$, $p = 0.018$), raw text Jaccard distance failed entirely to capture any statistically significant relationship ($r = 0.023$, $p = 0.812$). This contrast demonstrates that direct vocabulary overlap is too sparse to detect cultural lineage, whereas our motif representation successfully filters out noise to capture the underlying intercultural correlations.

**Religious motives.** To test our method on a new dataset, we ran it on our religious stories. Since this dataset is smaller, we did not cluster the motifs, but only merged a motif with the motifs of which is it a prefix. We discovered three "universal" motifs that appear across stories of all religions in our dataset: "Journey undertaken", "Magical helper aids", and "Reward earned". The motifs we find are not necessarily representative of the religions, because they depend heavily on our dataset, which is small and probably biased. Still, the similarity between the motifs in the stories (Figure 5) corresponds to the shared heritage between the religions: The Indian religions (Buddhism, Hinduism, and Jainism) share many motifs, as do the Abrahamic ones (Judaism, Christianity, and Islam), the East-Asian ones (Chinese Folk Religion and Taoism), and the Ancient Roman and Greek ones (which additionally share links with the Abrahamic faiths). Shintoism, as an insular religion, sits markedly apart.

The weak link between Shintoism and Jainism is almost certainly spurious. The weak link between the Greek and Hindu traditions, however, could be interpreted as a hint at their common Proto-Indo-European heritage (Noyer, 2011; Mallory and Adams, 2006; Suchanek, 2026).

## 6 Conclusion

In this paper, we have shown how motifs can be extracted automatically from stories, tales, and religious legends. We have found that these motifs capture the main events of the story, and overlap largely with the manually extracted motifs from domain experts. Our motifs can be used to mitigate gender bias in existing motif collections, to visualize and confirm hypotheses about the geographic origin of stories, and to analyze common story flows. They can also serve to study the relatedness of religious traditions.

Our datasets are small and possibly biased, and thus our results should not be taken as authoritative findings about cultures, languages, or religions. They are merely indicators that could inspire domain experts to find and validate hypotheses. We hope that our work can help extract motifs from other sources, and aid with the discovery of narrative patterns in and across cultures. Future work could cover not just motifs, but also concepts that appear in stories (such as "apples", "keys", or "dragons") and their distribution across cultures (Nakawake and Sato, 2019).

All our data and code is available at `https://github.com/sabashv/folktale-analysis`, as far as the licenses of our sources permit.

**Acknowledgements.** This work was partially supported by the Hi! PARIS and ANR France 2030 program (ANR-23-IACL-0005).

## 7 Limitations

While our datasets are the largest ones that are publicly available, they remain selective. In particular, they exhibit a heavy Indo-European focus. Our results are therefore not representative. Furthermore, we often do not have a ground truth for evaluation. Our results would thus have to be validated by domain experts.

## 8 Use of Generative AI

Generative AI was used to help polish the text of the manuscript.

## References

Antti Aarne. 1910. *Verzeichnis der Märchentypen*, volume 3 of *FF Communications*. Suomalaisen Tiedeakatemian Toimituksia, Helsinki.

A. Acharya, D. C. Estrada, S. Dahal, W. V. H. Yarlott, D. Gomez, and M. A. Finlayson. 2024. Discovering implicit meanings of cultural motifs from text. In *Proceedings of the Sixth Workshop on Natural Language Processing and Computational Social Science (NLP+CSS)*, pages 46–56.

D. L. Ashliman. 2021. Folktexts: A library of folktales, folk legends, fairy tales, and mythology. Accessed: 2025-05-17.

Giambattista Basile. 1634–1636. *Lo cunto de li cunti overo lo trattenemiento de peccerille*. Stamperia di Giovanni Battista Gargano e Luigi Nucci, Napoli. Part 1 published 1634, Part 2 published 1636; often collectively known as Il Pentamerone.

Yuri Berezkin and Evgeny Duvakin. 2016a. Buried in a head: African and asian parallels to aesop's fable. *Folklore*, 127(1):91–102.

Yuri Berezkin and Evgeny Duvakin. 2016b. Buried in a head: African and asian parallels to aesop's fable. *Folklore*, 127(1):91–102.

Sara Graça Da Silva and Jamshid J Tehrani. 2016. Comparative phylogenetic analyses uncover the ancient roots of indo-european folktales. *Royal Society open science*, 3(1):150645.

Scott Deerwester, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman. 1990. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):391–407.

Alessandra Flaccavento, Youri Peskine, Paolo Papotti, Riccardo Torlone, and Raphael Troncy. 2025. Automated detection of tropes in short texts. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 5936–5951, Abu Dhabi, UAE. Association for Computational Linguistics.

Jonathan Gottschall. 2012. *The Storytelling Animal: How Stories Make Us Human*. Houghton Mifflin Harcourt, Boston.

Jacob Grimm and Wilhelm Grimm. 1812. *Kinder- und Hausmärchen*. P. Hamm.

J. Hagedorn and S. Darányi. 2022. Bearing a bag-of-tales: An open corpus of annotated folktales for reproducible research. *Journal of Open Humanities Data*, 8(16):1–10.

Suzanne Keen. 2013. Narrative empathy. In Peter Hühn et al., editors, *The Living Handbook of Narratology*. Hamburg University, Hamburg.

Jeanne-Marie Leprince de Beaumont. 1798. *Magasin des enfans, ou Dialogues d'une sage gouvernante avec ses élèves de la première distinction*, volume 2. Veuve Rusand, Lyon.

J. P. Mallory and D. Q. Adams. 2006. *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford University Press.

Nathan Mantel. 1967. The detection of disease clustering and a generalized regression approach. *Cancer research*, 27(2_Part_1).

Yo Nakawake and Kosuke Sato. 2019. Systematic quantitative analyses reveal the folk-zoological knowledge embedded in folktales. *Palgrave Communications*, 5(1).

Dong Nguyen, Dolf Trieschnigg, Theo Meder, and Mariët Theune. 2012. Automatic classification of folk narrative genres. In *11th Conference on Natural Language Processing, KONVENS 2012, Empirical Methods in Natural Language Processing, Vienna, Austria, September 19-21, 2012*, pages 378–382.

Dong Nguyen, Dolf Trieschnigg, and Mariet Theune. 2013. Folktale classification using learning to rank. In *Advances in information retrieval. Proceedings*, pages 195–206.

Rolf Noyer. 2011. Proto-indo-european language and society: Pie dieties and the sacred. Lecture notes/slides.

Charles Perrault. 1697. *Histoires ou contes du temps passé, avec des moralités: Contes de ma mère l'Oye*. Barbin, Paris.

Steven Pinker. 2012. *The Better Angels of Our Nature: Why Violence Has Declined*. Penguin Books.

Vladimir Propp. 1968. *Morphology of the Folktale*. University of Texas Press, Austin.

Giulia Pucci, Fabio Massimo Zanzotto, and Leonardo Ranaldi. 2025. Animate, or inanimate, that is the question for large language models. *Information*, 16(6):493.

Katrin Pöge-Alder. 2007. *Märchenforschung. Theorien, Methoden, Interpretationen*. Gunter Narr Verlag, Tübingen.

Barkha Singh. 2023. Women folklore: The unofficial custodian of culture. *International Journal of Education, Modern Management, Applied Science & Social Science (IJEMMASSS)*, 5(2):134–138.

Fabian M. Suchanek. 2026. *The Atheist Bible — A Secular Humanist view on life, the Universe, and everything*. Kindle Direct Publishing, Paris. Chapter 11.

Stith Thompson. 1955–1958. *Motif-Index of Folk-Literature: A Classification of Narrative Elements in Folktales, Ballads, Myths, Fables, Mediaeval Romances, Exempla, Fabliaux, Jest-Books and Local Legends*, volume 1–6. Indiana University Press, Bloomington.

Stith Thompson. 1961. *The Types of Folktale: A Classification and Bibliography*, 2nd revision edition. Academia Scientiarum Fennica, Helsinki.

Marc Thuillard, Jean-Loïc Le Quellec, Julien d'Huy, and Yuri Berezkin. 2018. A large-scale study of world myths. *Trames: A Journal of the Humanities and Social Sciences*, 22(4):A1–A44.

H.J. Uther. 2011. *The Types of International Folktales: A Classification and Bibliography, Based on the System of Antti Aarne and Stith Thompson*. FF communications. Suomalainen Tiedeakatemia, Academia Scientiarum Fennica.

Yashpeh. 2013. Yashpeh: International folktales collection. `https://folkmasa.org/yashpeh/yashpeh.htm`. Accessed: 2026-07-30.

Gaifan Zhang, Yi Zhou, and Danushka Bollegala. 2024. Evaluating unsupervised dimensionality reduction methods for pretrained sentence embeddings. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*.