

How People Describe Their Mystical Experiences: Analyzing Language, Narrative, and Digital Expression

Daniel Zhu *

St Andrew's School, Middletown, Delaware, The United States of America

* Corresponding Author Email: danielzhuforcollege@gmail.com

Abstract. This study investigates how mystical experiences are articulated in digital environments, especially in religion-focused communities (“subreddits”) on Reddit, a popular social media platform. Classical theorists like William James argued that mystical experiences resist linguistic expression; in other words, according to the traditional scholarly view, these phenomena are inherently “ineffable.” However, this study finds that social media users share detailed narratives of their mystical experiences through a blend of metaphor, emotion, and storytelling. Using the Reddit API, data are retrieved from five subreddits representing different faiths and religious emphases: r/religion, r/Christianity, r/spirituality, r/Buddhism, and r/DebateAnAtheist. Mixed-method analyses were then conducted, combining computational text analysis, including VADER sentiment analysis and Latent Dirichlet Allocation (LDA) topic modeling, with qualitative techniques, including narrative and cultural analysis. In addition to the abundance of effable expression of mystical experience, it is also revealed that community identity significantly shapes how mystical experiences are framed. For instance, r/Buddhism emphasizes meditative mechanics and doctrine; r/Christianity mostly relies on language derived from the Scripture and the Church; r/DebateAnAtheist focuses almost exclusively on epistemology and materialist inquiry with little experiential description. Thus, the study argues that “ineffability” may be a product of social environment rather than an inherent quality of the experience. When digital anonymity and community-specific “scripts” provide individuals an opportunity to escape the confines of the traditional social environment, their subjective religious encounters are translated into detailed, though socially negotiated, narratives. This study demonstrates that digital environments, such as social media, are a crucial platform where individuals can seek meaning, validation, and identity.

Keywords: Mystical experiences; Digital religion; Latent Dirichlet Allocation (LDA); VADER sentiment analysis.

1. Introduction

Mystical experiences have long been avoided as a topic of discussion due to their inherent “ineffability”. As a result, individuals often struggle to articulate these experiences using conventional language. Classical scholars such as William James [1] and Walter Stace [2] argued that mystical experiences are fundamentally beyond linguistic expression. Later scholars have challenged this view, suggesting that such experiences are always shaped by cultural and linguistic frameworks rather than existing as purely ineffable phenomena [3, 4].

However, the rise of online platforms has created new spaces where users feel more comfortable describing and reflecting on deeply personal phenomena. Online communities such as Quora, Discord, and Reddit have become increasingly popular in recent years due to their well-designed platforms for open discussion among anonymous users. Specifically, Reddit hosts numerous communities focused on religion and theology. These sophisticated “subreddits” contain valuable information on how online users express their ideas and perspectives on faith and belief systems. Crucially, the nature of anonymity lowers social risk and reduces fear of judgment, encouraging users to share vulnerable, unconventional, or highly subjective experiences they might otherwise withhold, a phenomenon described as the “online disinhibition effect” [5]. In addition, scholars of digital religion argue that online environments do not merely host religious expression but actively shape how beliefs and experiences are communicated and understood [6].

Despite established claims that mystical experiences resist language altogether, Reddit users frequently share long, detailed narratives that fervently attempt to capture these moments through metaphor, emotion, and storytelling. This contrast raises paramount questions about how language and narrative structure shape the expression of experiences traditionally deemed indescribable. Prior research suggests that individuals rely on culturally available narratives and interpretive frameworks to make sense of extraordinary experiences [7, 8]. Accordingly, this study analyzes how mystical experiences are narrated in digital environments, comparing linguistic, emotional, and thematic patterns across religious and non-religious subreddits. By examining the role of online platforms in shaping collective interpretation and identity formation, this research revisits the theoretical debate on ineffability, introduces a mixed-methods approach combining computational text analysis and narrative analysis, and highlights how contemporary individuals seek meaning, support, and self-understanding through digital storytelling.

Subsequently, this study proposes the following three research questions.

1. How do people from different religious / non-religious communities describe mystical experiences in online discussions?
2. What emotional tones characterize these narratives, and how do they vary across communities?
3. What underlying themes emerge from these discussions, and how are they distributed among different subreddit cultures?

This study is significant on several levels. Theoretically, it offers a renewed examination of long-standing scholarly claims about the ineffability of mystical experiences by investigating how such experiences are articulated in contemporary digital contexts. Particularly, empirical insight is provided from an analysis of large-scale, naturally occurring narratives on subreddits. Methodologically, this study bridges computational linguistics and the study of religion by combining natural language processing techniques (such as sentiment analysis and topic modeling) with close narrative and cultural analysis of online texts. Socially, this research sheds light on how contemporary individuals seek meaning, validation, and identity through online storytelling, revealing the role of digital platforms in transforming private, subjective experiences into shared, socially negotiated narratives.

2. Data and Methods

2.1. Data Collection and Preprocessing

Data for this study were collected from the Reddit API using Python in the Spyder environment. We targeted multiple relevant subreddits: *r/religion*, *r/Christianity*, *r/spirituality*, *r/Buddhism*, and *r/DebateAnAtheist*, chosen for their distinct religious or spiritual orientations and the abundance of threads about mystical experiences contained. To capture relevant content comprehensively, we searched these subreddits using the keyword “mystical experiences”. Ultimately, we retrieved up to 669 of the most recent submissions (posts) across all subreddits analyzed.

Importantly, for each retrieved submission, we also collected all associated comments (including nested replies) to preserve the full context of the discussion and all relevant information. Consequently, the results were organized into two datasets: one containing submission-level information (title, selftext, creation time, etc.) and another containing comment- or reply-level information (comment text, creation time, etc.). This dual structure preserves the hierarchical nature of Reddit conversations while enabling flexible analytical strategies.

After all, this data collection method has no underlying ethical concerns. First, data was retrieved via the official Reddit API, ensuring compliance with the platform's protocols. Second, Reddit posts are situated within the public domain; by adhering to the platform's terms of service, users acknowledge that their contributions are accessible for external use. Third, supplementing the anonymity inherent to Reddit, this study excludes all personally identifiable information from the analysis. This

methodology aligns with standard ethical guidelines for analyzing public online content while prioritizing the privacy of individual contributors.

We used rigorous preprocessing methods to prepare the collected raw textual data for analysis. Non-informative or low-value elements such as URLs, emojis, Reddit-specific markup (e.g., /u/, /r/, [deleted]), and formatting artifacts were removed to reduce noise. Subsequently, we applied standard natural language processing steps: all text was converted to lowercase, punctuation and stop words were removed, and the texts were tokenized. Furthermore, posts and comments with extremely low content length or irrelevant content were filtered out. To maintain semantic richness, lemmatization rather than stemming was used, as many users employed metaphorical and poetic language to describe mystical experiences. The cleaned, lemmatized texts from both submissions and comments were then optionally combined into a single corpus for unified analysis, enabling a holistic view of mystical discourse across different religious communities

2.2. Data Analysis

2.2.1. Lexical and Narrative Exploration

The first phase of analysis examined how mystical experiences are discussed across different religious communities, focusing on the lexical patterns and narrative structures of online discourse. For each subreddit, we summarized the dataset by the number of total posts and comments, average text length, and the typical depth of comment threads, providing a descriptive overview of the discussion landscape.

Lexical exploration then placed greater emphasis on vocabulary related to religious and experiential content than on general functional words. For this purpose, we identified both faith-specific references (e.g., Jesus, church, karma, nirvana, and meditation) and descriptive words that captured how users articulated their mystical experiences, including adjectives and concrete nouns such as light, vision, presence, dream, and awakening. To highlight the distinctive linguistic features of each community, we computed TF-IDF (Term Frequency-Inverse Document Frequency) scores and generated comparative tables of the most characteristic words per subreddit. This approach marginalized common discourse markers—such as 'think' or 'say'—to better clarify the unique lexical signatures inherent to different religious contexts."

Finally, to better understand the narrative form of these discussions, we calculated simple structural indicators, including average sentence length, the frequency of first-person pronouns, and the presence of temporal or spatial markers (e.g., yesterday, in church, during meditation). These metrics substantiate the qualitative observation that users typically prioritize the description of sensations or insights over the construction of complete experiential narratives. The results of this phase—comprising subreddit-level summaries, TF-IDF keyword comparisons, and targeted word clouds—establish a descriptive baseline for the thematic and emotional analyses that follow.

Table 1. Descriptive Statistics of the Reddit Dataset by Subreddit and Content Type

Subreddit	Category	Numbers (N)	Average Comments + replies/posts
r/religion	Submission (posts)	96	14
	Comments (Replies)	1332	
r/Christianity	Submission (posts)	200	12
	Comments	2470	
r/Spirituality	Submission (posts)	198	7
	Comments (replies)	1413	
r/Buddism	Submission (posts)	132	15
	Comments (replies)	2073	
r/DebateAnAtheist	Submission (posts)	43	136
	Comments (Replies)	5833	

Table 2. Categories of Three Types of Keyword Dictionaries

Broad Category	Sub-category	Word List
Religious words	Christianity	"jesus", "christ", "christian", "church", "bible", "god", "lord", "prayer", "heaven", "hell", "angel", "holy", "spirit", "trinity", "cross", "pastor", "psalm", "gospel", "salvation"
	Buddism	"buddha", "buddhist", "dharma", "karma", "nirvana", "sutra", "temple", "meditation", "enlightenment", "monk", "lotus", "sangha"
	Hinduism / Eastern religions	"shiva", "krishna", "hindu", "hinduism", "yoga", "ashram", "guru", "mantra", "veda", "chakra"
	General spirituality	"spiritual", "faith", "divine", "ritual", "sacred", "oracle", "pilgrimage", "soul", "afterlife", "miracle"
Experimental words	Sensory experiences	"light", "vision", "voice", "sound", "music", "smell", "touch", "warmth", "cold", "vibration",
	Emotional/psychological states	"Peace", "calm", "fear", "joy", "bliss", "ecstasy", "awe", "wonder", "love", "compassion", "emptiness", "clarity", "confusion", "shock", "surprise",
	Transcendental/mystical imagery	"presence", "spirit", "dream", "floating", "flying", "out of body", "transcendence", "infinite", "radiance", "glow", "mystical", "sacred", "holy", "divine", "energy", "force"
General words	Common pronouns	"i", "me", "my", "we", "you", "they", "he", "she", "it", "one", "someone"
	Common verbs	"think", "say", "know", "feel", "see", "believe", "understand", "guess", "talk", "ask", "wonder", "hear"
	Discourse/logic markers	"maybe", "probably", "actually", "basically", "like", "just", "really", "well", "so", "then", "if", "but", "because"

Table 3. Average Word Category Frequency Ratios by Subreddit

Subreddit	religious_ratio	experiential_ratio	general_ratio
r/religion	0.041109	0.022526	0.044441
r/Christianity	0.074188	0.023220	0.045628
r/Spirituality	0.026669	0.027257	0.042419
r/Buddism	0.031940	0.014786	0.046272
r/DebateAnAtheist	0.022673	0.009932	0.055308

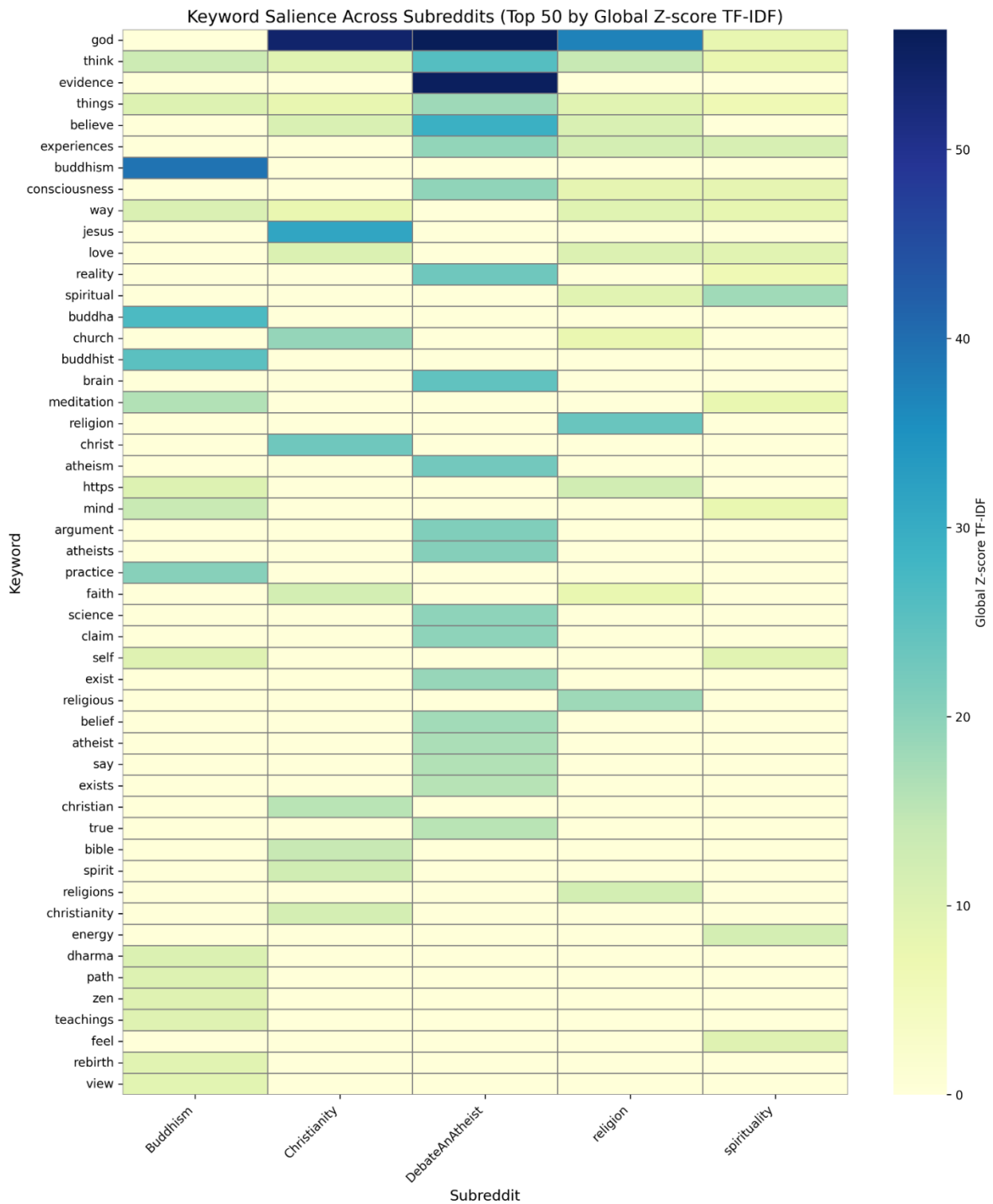


Figure 1. Keyword Saliency Across Subreddits (Top 50 by Global Z-score TF-IDF)

Table 4. Narrative Features by Subreddit

Subreddit	First_person_perspectives	Temporal Markers	Distribution of part-of-speech ratios			
			Noun	Verb	Adjective	Adverb
r/religion	0.000203	2.791667	0.450451	0.200450	0.213786	0.071010
r/Christianity	0.000273	3.735000	0.438557	0.222541	0.205809	0.073603
r/Spirituality	0.000267	5.232323	0.436100	0.219740	0.208561	0.076769
r/Buddhism	0.000300	2.651515	0.435446	0.214695	0.203087	0.081584
r/DebateAnAtheist	0.000130	1.860465	0.555862	0.155603	0.137275	0.084787

2.3. Sentiment Analysis

The second stage of our analysis utilizes the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analyzer to quantify the emotional tone of the collected data. VADER assigns each post, comment, or reply a negative score within the range of $[-1, 0)$ and a positive score within $(0, 1]$. This metric allows for the calculation of composite scores for both individual threads and entire subreddits, representing the aggregate emotional valence expressed by users.

We then isolated threads with high emotional intensity (positive > 0.8 ; negative < -0.5) to identify frequently used 'extreme words.' For instance, a thread scoring -0.6 regarding disbelief in the afterlife often contains terms like 'death,' 'hell,' or 'nonsense.' Tracking these patterns helps demonstrate how specific vocabulary drives the emotional tone of the discussion. By analyzing this data, we identified consistent word patterns linked to emotional extremes. This reveals how specific terms and themes correlate with strongly positive or negative sentiments across different subreddits.

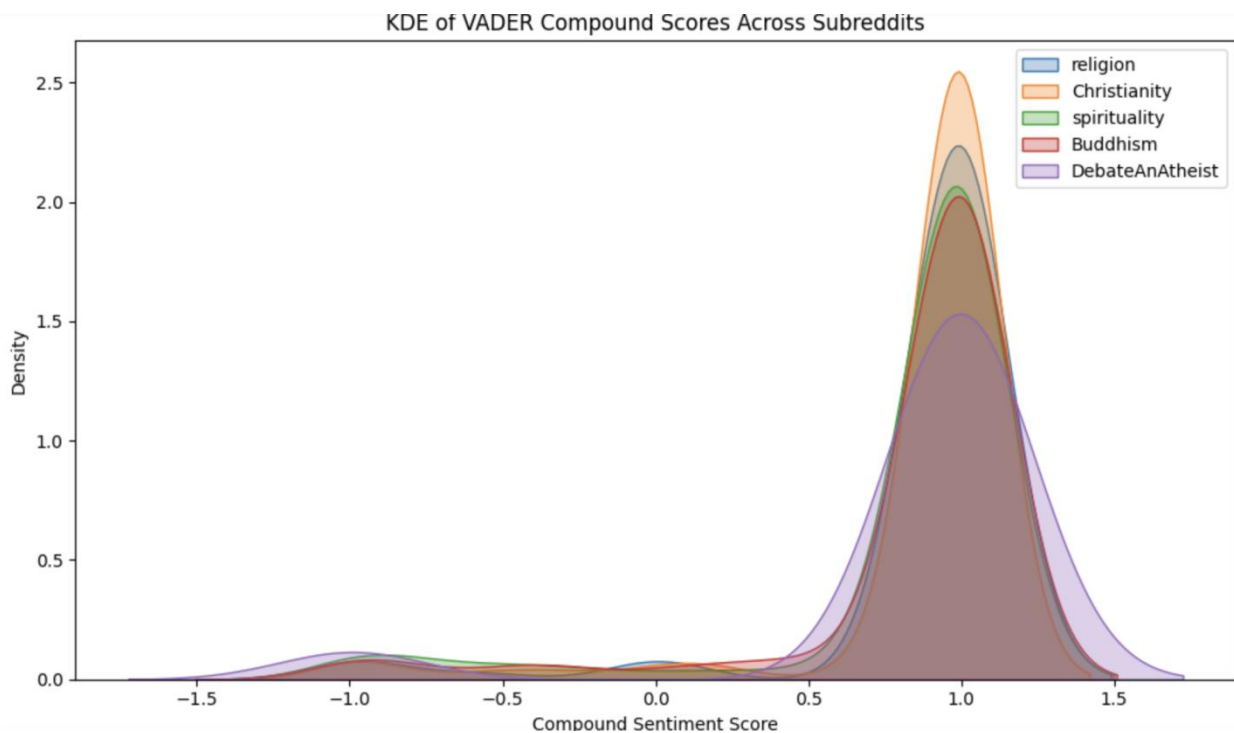


Figure 2. KDE of VADER Compound Score Across Subreddits

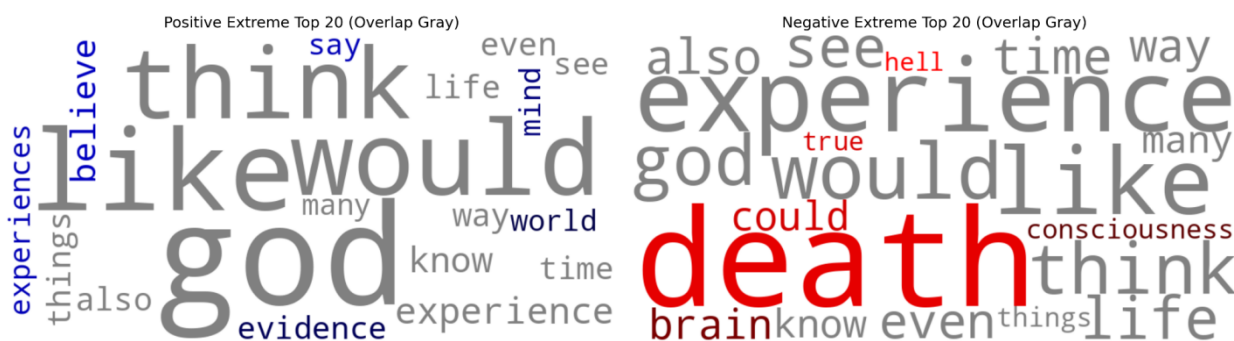


Figure 3. Word Cloud Visualization of Sentiment Extremes in Mystical Experience Narratives.

2.4. Thematic Analysis / Topical Modeling

To examine the latent thematic structures underlying users' narratives of mystical experiences, we employed Latent Dirichlet Allocation (LDA) for topic modeling. Topic modeling is a computational method that identifies recurring patterns of words across a large collection of text, effectively uncovering the "hidden" themes that define different conversations. We conducted this analysis in two stages.

First, we evaluated candidate models with topic numbers (K) ranging from 4 to 10, utilizing CV coherence scores to measure semantic consistency. Higher coherence scores were prioritized as they yield topics that are more interpretable and aligned with human judgment. Based on the highest coherence trends and manual inspection for interpretability, we selected a six-topic model (K = 6) for the main analysis. In the second stage, each corpus was represented as a distribution over these topics and aggregated by subreddit to identify how different communities emphasize distinct narrative patterns. This approach enabled a systematic comparison across religious, spiritual, and secular spaces, allowing us to link individual accounts to broader identity frameworks without imposing predefined thematic categories.

2.5. Thematic Details of the Six-Topic Model

The Latent Dirichlet Allocation (LDA) analysis revealed six distinct thematic clusters that characterize mystical discourse across the subreddits:

Topic 1: ['god', 'jesus', 'christ', 'church', 'spirit', 'christian', 'bible', 'love', 'holy', 'father']

Christian Theology: centered on institutional and doctrinal terms. This topic reflects traditional religious interpretations, in which experiences are framed within established Christian frameworks.

Topic 2: ['spiritual', 'meditation', 'experiences', 'mind', 'feel', 'felt', 'way', 'see', 'body', 'get']

Somatic and Meditative Practice: focuses on the physical and psychological process of the experience. It suggests a focus on the mechanics of achieving or processing mystical states.

Topic 3: ['god', 'believe', 'think', 'evidence', 'things', 'religion', 'even', 'belief', 'say', 'reality']

Epistemology and Belief: captures the rhetorical struggle over the validity of mystical claims. It shifts from describing the experience to the intellectual process of justifying knowledge, weighing evidence, and negotiating truth within a debate-oriented framework.

Topic 4: ['consciousness', 'brain', 'think', 'evidence', 'reality', 'science', 'mind', 'way', 'even', 'conscious']

Materialist and Scientific Inquiry: represents secular or skeptical attempts to explain mystical phenomena through neuroscience or cognitive science.

Topic 5: ['buddhism', 'think', 'buddhist', 'buddha', 'many', 'good', 'way', 'religion', 'much', 'even']

Buddhist Philosophy: reflects a narrative style rooted in Buddhist philosophical frameworks. It emphasizes the use of specialized community terminology and doctrinal concepts to structure and validate mystical experiences within this specific tradition.

Topic 6: ['god', 'soul', 'world', 'spiritual', 'love', 'self', 'divine', 'heaven', 'mind', 'death']

Metaphysical and Existential Soul: represents a broad spiritual framework focused on ultimate meaning. It transcends institutional boundaries to explore the nature of the self and the divine, reflecting a narrative concern with life, death, and the universal essence of mystical encounters.

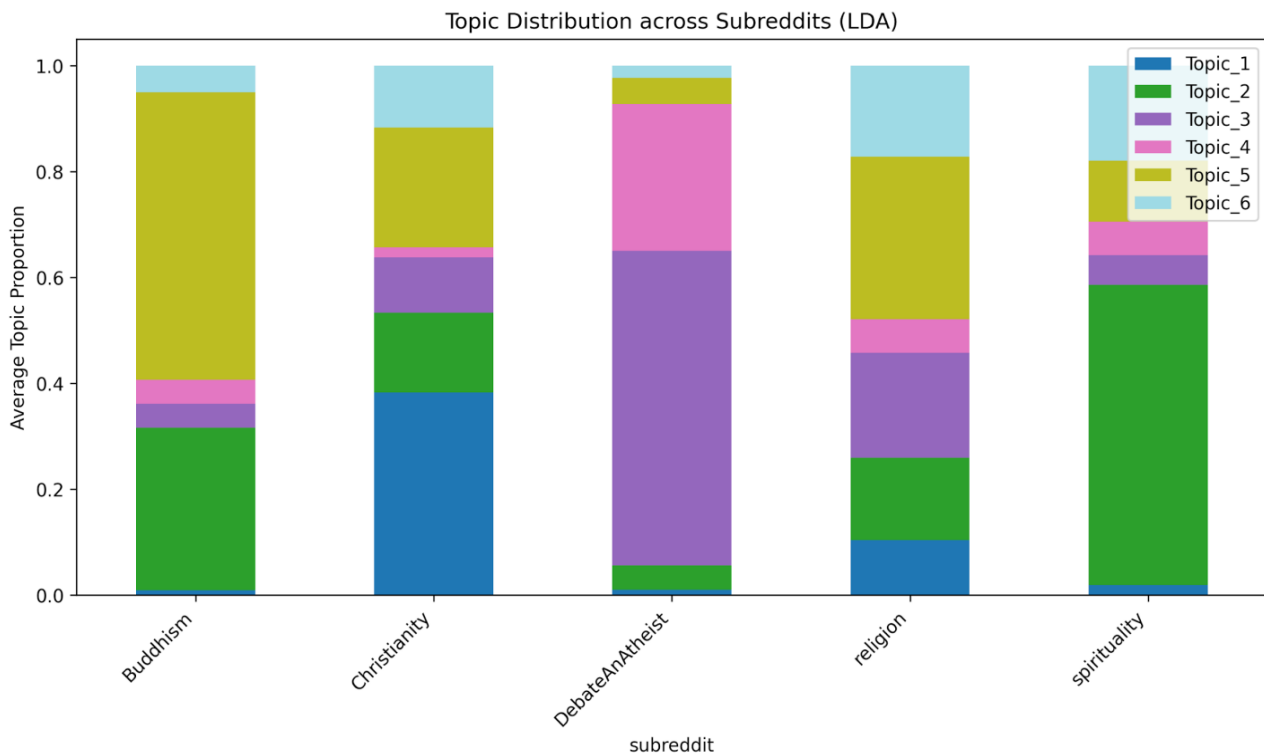


Figure 4. Average Topic Distribution across Sampled Subreddits via Latent Dirichlet Allocation (LDA)

3. Distribution and Interpretation

This distribution graph shows how the six topics are spread across the five target subreddits, confirming that community identity significantly shapes the narrative framing of mystical experiences.

Dominance of Tradition-Specific Frameworks:

- a. {r/Buddhism} is overwhelmingly dominated by Topic 5 (Buddhist Philosophy) and Topic 2 (Somatic/Meditative Practice). This follows logic, suggesting that users in this community describe experiences primarily through specific doctrinal lenses and the physical mechanics of meditation.
- b. {r/Christianity} shows a unique concentration in Topic 1 (Christian Theology), which is nearly absent in other subreddits. This indicates a heavy reliance on institutional and scriptural language to interpret mystical events.

The Secular-Skeptical Divide:

- a. {r/DebateAnAtheist} is almost entirely composed of Topic 3 (Epistemology and Belief) and Topic 4 (Materialist Inquiry). Unlike the other groups, this community largely ignores experiential or theological narratives in favor of debates over evidence, logic, and scientific explanations.

Spiritual and General Pluralism:

- a. {r/Spirituality} and {r/religion} display the most diverse distributions. While r/Spirituality leans heavily into Topic 2 (Practice) and Topic 6 (Metaphysical Soul), {r/religion} includes significant portions from all topics.

The graph provides empirical evidence that mystical experiences are socially negotiated. Rather than a single "ineffable" experience, the data shows that users select specific "scripts"—whether theological, scientific, or meditative—based on the cultural logic of their chosen digital community.

4. Discussions and Conclusions

The results of this study directly challenge long-standing scholarly claims regarding the "ineffability" of mystical experiences. Classical theorists have historically argued that such states resist linguistic expression and exist beyond the reach of conventional vocabulary [1, 2]. However, the data from these five subreddits reveals a different reality: contemporary individuals are not only willing but highly capable of describing these encounters in rich, detailed, and articulate language.

Rather than remaining "indescribable," mystical experiences on Reddit are characterized by high frequencies of sensory markers, emotional nuances, and vivid metaphors. This finding aligns with constructivist perspectives in the sociology of religion, which suggest that experiences are shaped and structured through available linguistic and cultural repertoires [3,4]. It also supports the argument that emotional expression itself is socially organized and practiced, rather than purely spontaneous [9].

This suggests that "ineffability" may not be an inherent property of the experience itself, but rather a reflection of the social environments in which they are shared. In public or institutional contexts, where religious claims often carry social risk, skepticism, or theological disagreement, individuals may hesitate to articulate mystical experiences openly or may lack a receptive audience willing to take such claims seriously. In digital spaces, the combination of anonymity and community-specific narrative "scripts" enables users to translate intensely subjective experiences into forms of language that can be publicly articulated and understood [5, 6]. More broadly, these findings suggest that mystical experiences are not simply private, ineffable events, but are actively constructed through narrative practices that allow individuals to interpret, communicate, and legitimize their experiences within specific social contexts [7].

References

- [1] James, W. (1902). *The varieties of religious experience*. Longmans, Green & Co.
- [2] Reddy, W. M. (2001). *The navigation of feeling: A framework for the history of emotions*. Cambridge University Press.
- [3] Katz, S. T. (1978). Language, epistemology, and mysticism. In S. T. Katz (Ed.), *Mysticism and philosophical analysis* (pp. 22–74). Oxford University Press.
- [4] Proudfoot, W. (1985). *Religious experience*. University of California Press.
- [5] Stace, W. T. (1960). *Mysticism and philosophy*. Macmillan.
- [6] Campbell, H. A. (2013). *Digital religion: Understanding religious practice in new media worlds*. Routledge.
- [7] Suler, J. (2004). The online disinhibition effect. *Cyber Psychology & Behavior*, 7(3), 321–326.
- [8] Luhrmann, T. M. (2012). *When God talks back: Understanding the American evangelical relationship with God*. Knopf.
- [9] Taves, A. (2009). *Religious experience reconsidered: A building-block approach to the study of religion and other special things*. Princeton University Press.