

Text Analysis Report: blockchain_technology_deepseek

1. Basic Statistics

Metric	Value
Word Count	222
Sentence Count	10
Character Count	1477
Unique Word Count	138
Avg Word Length (chars)	5.51
Avg Sentence Length (words)	22.20
Avg Sentence Length (chars)	147.70
Sentence Length Std Dev (words)	6.03
Lexical Diversity (TTR)	0.6216
MATTR (Window=50)	0.8061

2. Entropy & Complexity

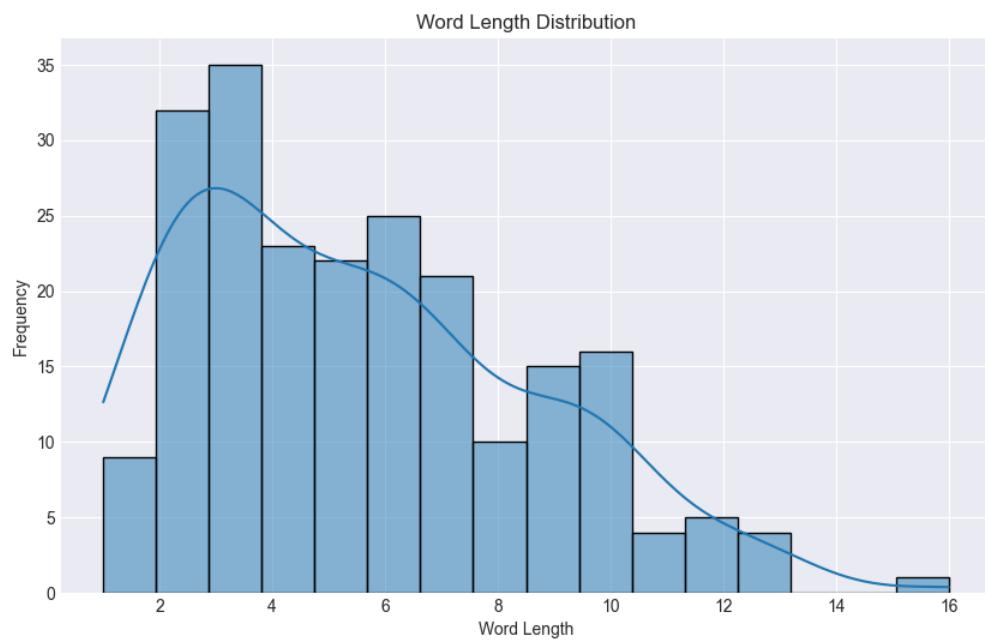
Metric	Value
Shannon Entropy (Words)	6.6565 bits
Shannon Entropy (Chars)	4.3871 bits
Function Word Ratio	0.3964

3. Readability Scores

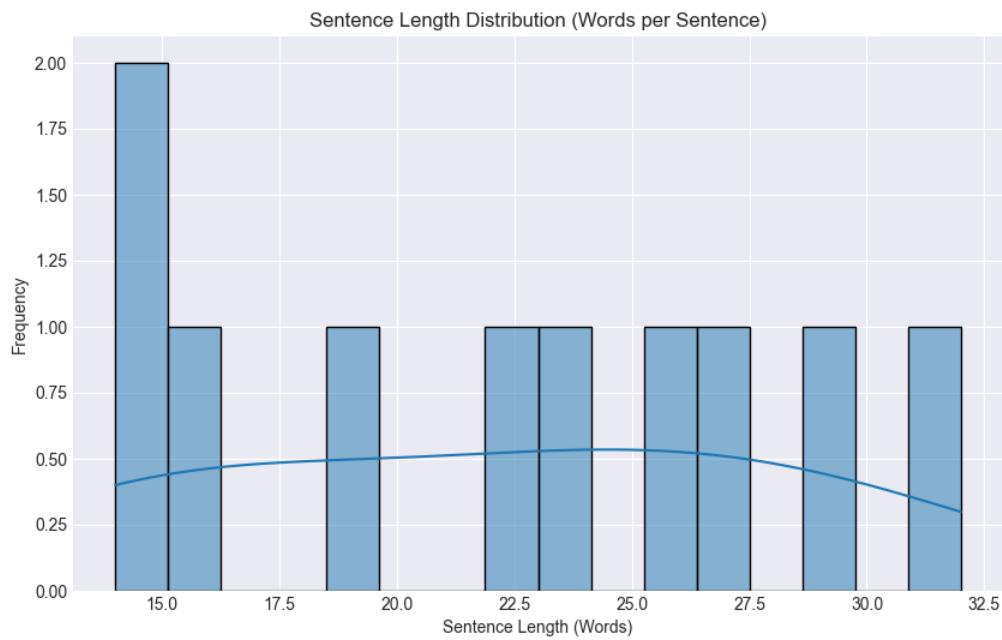
Index	Score
Flesch Reading Ease	32.33
Flesch-Kincaid Grade Level	14.20
Gunning Fog Index	16.25
SMOG Index	15.90
Coleman-Liau Index	15.14
Automated Readability Index	16.60

4. Visualizations

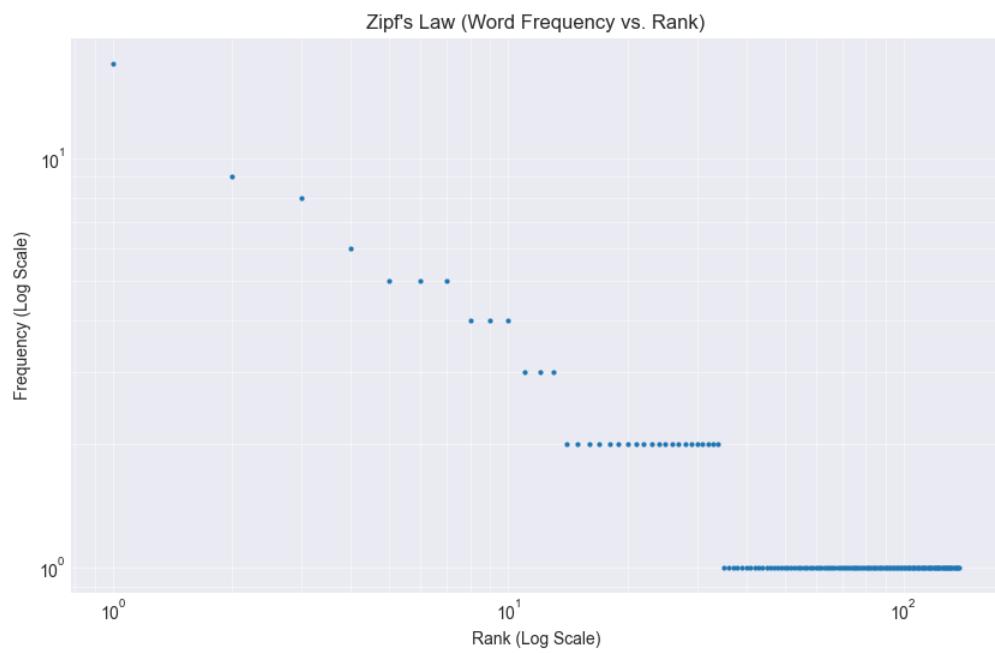
Word Length Distribution



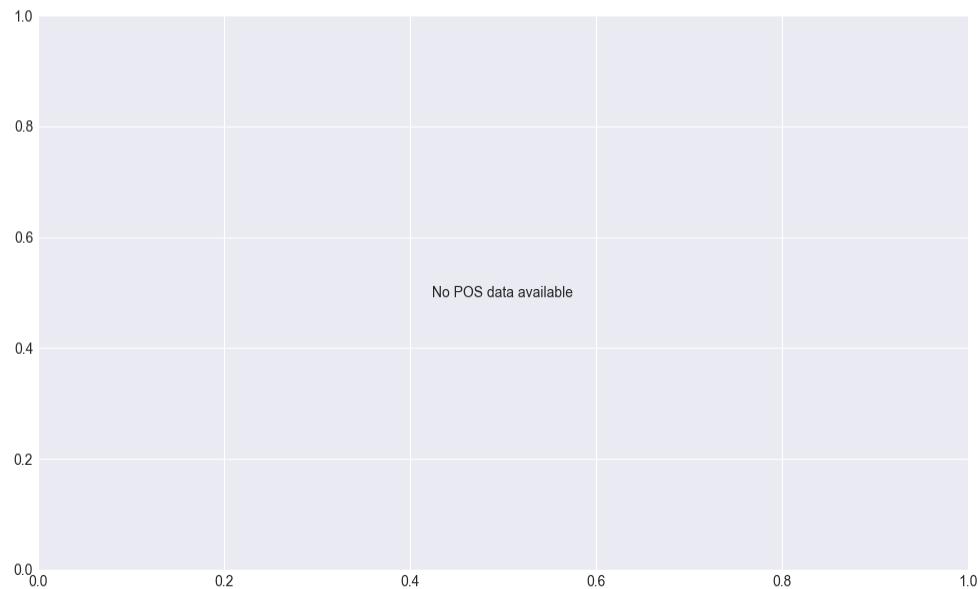
Sentence Length Distribution



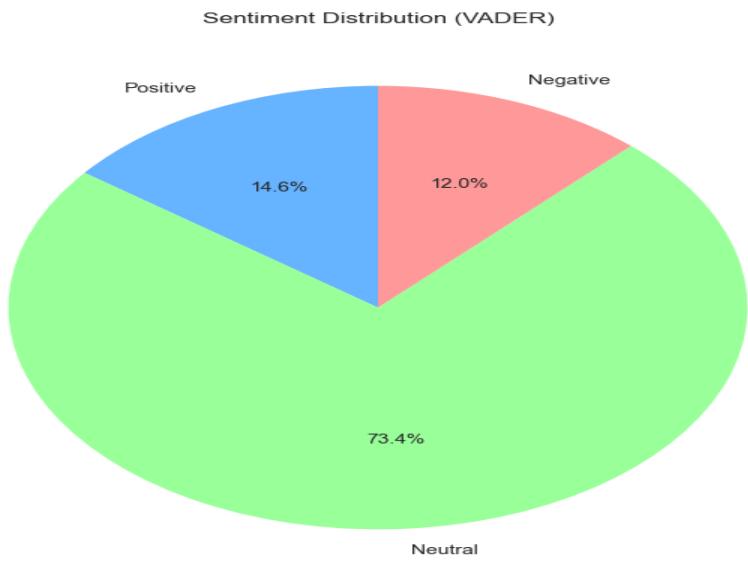
Zipf's Law Analysis (Word Frequency vs Rank)



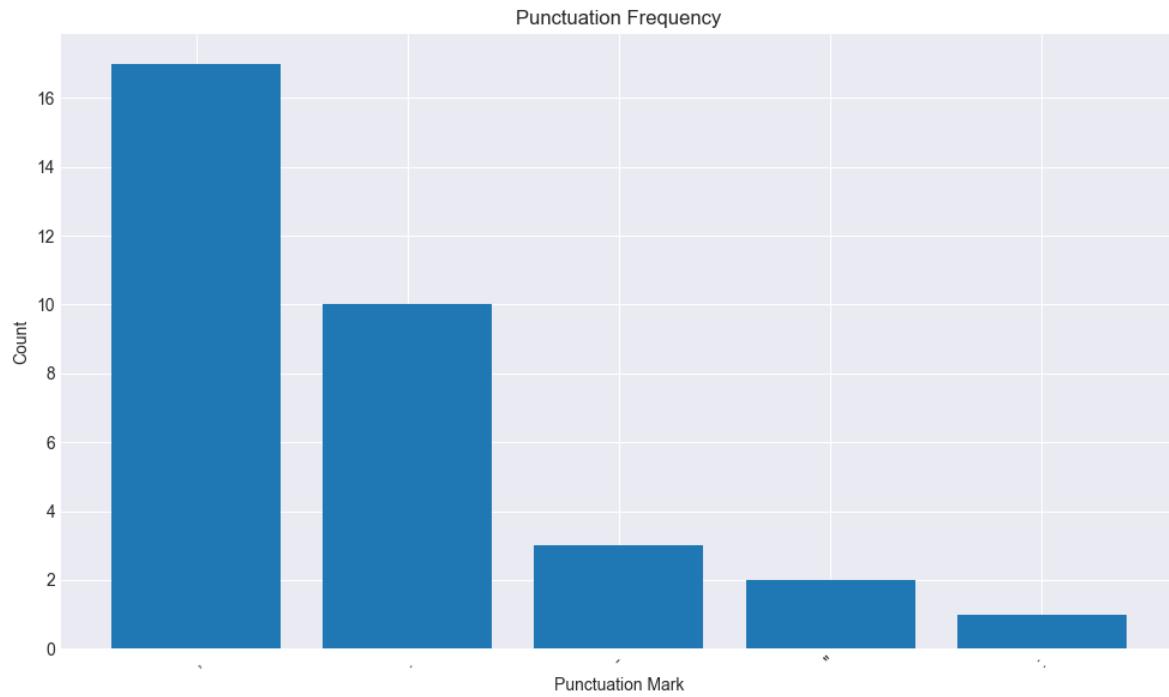
Part-of-Speech Distribution (Top 25)



Sentiment Distribution (VADER)



Punctuation Frequency



5. Frequency Analysis

Most Common Words (Top 15, with stopwords):

the (17), a (9), of (8), is (6), blockchain (5), and (5), to (5), technology (4), in (4), bitcoin (4), system (3), transactions (3), chain (3), decentralized (2), digital (2)

Most Common Punctuation:

',' (17), '.' (10), '-' (3), """ (2), ':' (1)

6. N-gram Analysis

Common 2-grams (Top 10):

'technology is' (2), 'is a' (2), 'it is' (2), 'the technology' (2), 'of the' (2), 'of blockchain' (2), 'blockchain was' (2), 'blockchain technology' (1), 'a decentralized' (1), 'decentralized digital' (1)

2-gram Repetition Rate: 0.0317

Common 3-grams (Top 10):

'blockchain technology is' (1), 'technology is a' (1), 'is a decentralized' (1), 'a decentralized digital' (1), 'decentralized digital ledger' (1), 'digital ledger system' (1), 'ledger system that' (1), 'system that records' (1), 'that records transactions' (1), 'records transactions across' (1)

3-gram Repetition Rate: 0.0000

Common 4-grams (Top 10):

'blockchain technology is a' (1), 'technology is a decentralized' (1), 'is a decentralized digital' (1), 'a decentralized digital ledger' (1), 'decentralized digital ledger system' (1), 'digital ledger system that' (1), 'ledger system that records' (1), 'system that records transactions' (1), 'that records transactions across' (1), 'records transactions across a' (1)

4-gram Repetition Rate: 0.0000

Character N-gram Analysis:

Common Character 3-grams (Top 10):

'the' (21), '_th' (20), 'he_' (17), 'in_' (16), 'tra' (11), 's_a' (10), 'ent' (10), 'ed_' (10), 'ing' (10), 'blo' (9)

Common Character 4-grams (Top 10):

'_the' (18), 'the_' (17), 'bloc' (9), 'lock' (9), 'ing_' (9), 'chai' (8), 'hain' (8), 'ain_' (8), '_of_' (8), '_blo' (8)

Common Character 5-grams (Top 10):

'_the_' (16), 'block' (9), 'chain' (8), 'hain_' (8), '_bloc' (8), 'lockc' (5), 'ockch' (5), 'ckcha' (5), 'kchai' (5), '_tran' (5)