

Text Analysis Report: blockchain_technology_chatgpt

1. Basic Statistics

| Metric | Value |
|---------------------------------|--------|
| Word Count | 130 |
| Sentence Count | 7 |
| Character Count | 987 |
| Unique Word Count | 90 |
| Avg Word Length (chars) | 6.45 |
| Avg Sentence Length (words) | 18.57 |
| Avg Sentence Length (chars) | 141.00 |
| Sentence Length Std Dev (words) | 6.43 |
| Lexical Diversity (TTR) | 0.6923 |
| MATTR (Window=50) | 0.8617 |

2. Entropy & Complexity

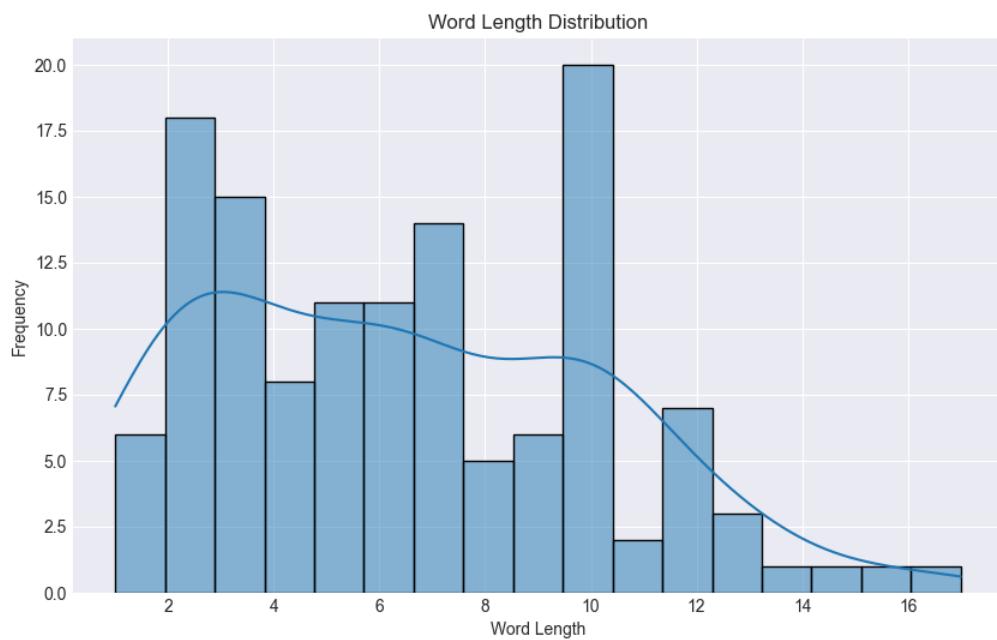
| Metric | Value |
|-------------------------|-------------|
| Shannon Entropy (Words) | 6.2425 bits |
| Shannon Entropy (Chars) | 4.4142 bits |
| Function Word Ratio | 0.3231 |

3. Readability Scores

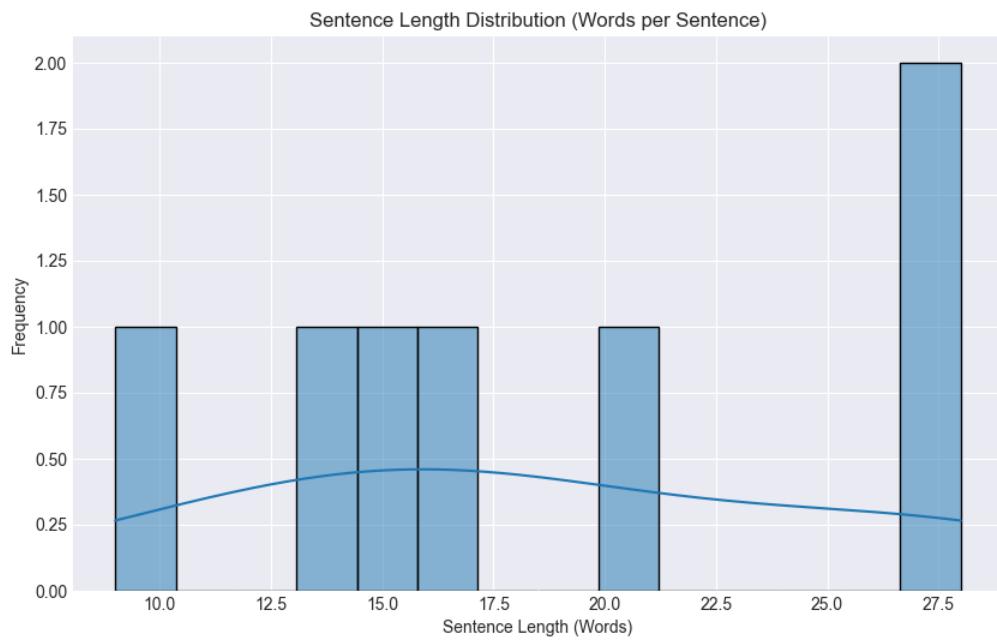
| Index | Score |
|-----------------------------|-------|
| Flesch Reading Ease | 10.50 |
| Flesch-Kincaid Grade Level | 16.40 |
| Gunning Fog Index | 17.90 |
| SMOG Index | 17.50 |
| Coleman-Liau Index | 20.42 |
| Automated Readability Index | 19.10 |

4. Visualizations

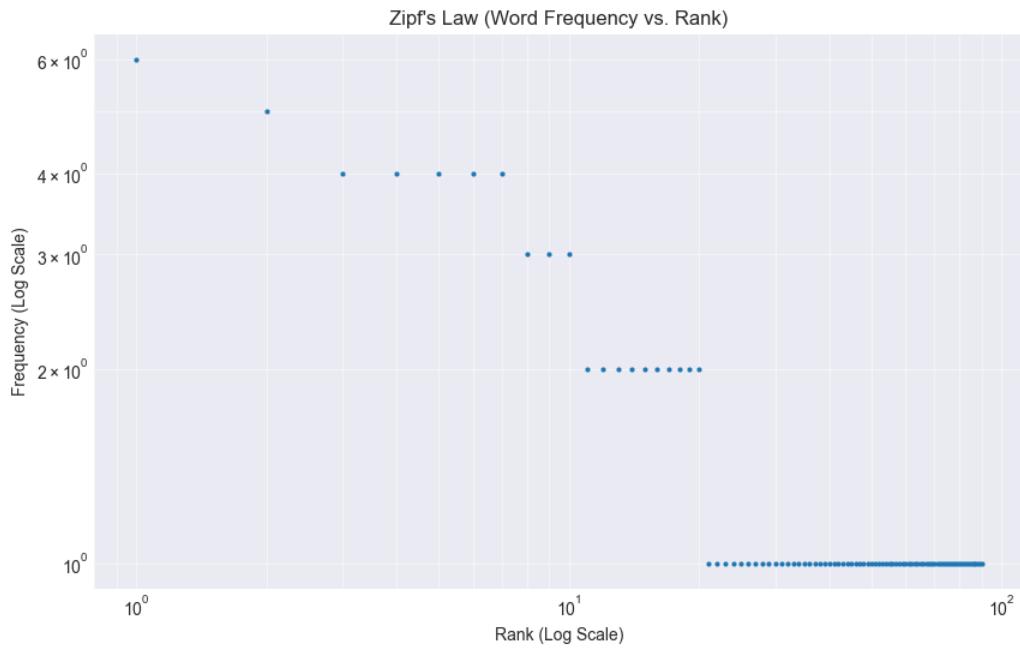
Word Length Distribution



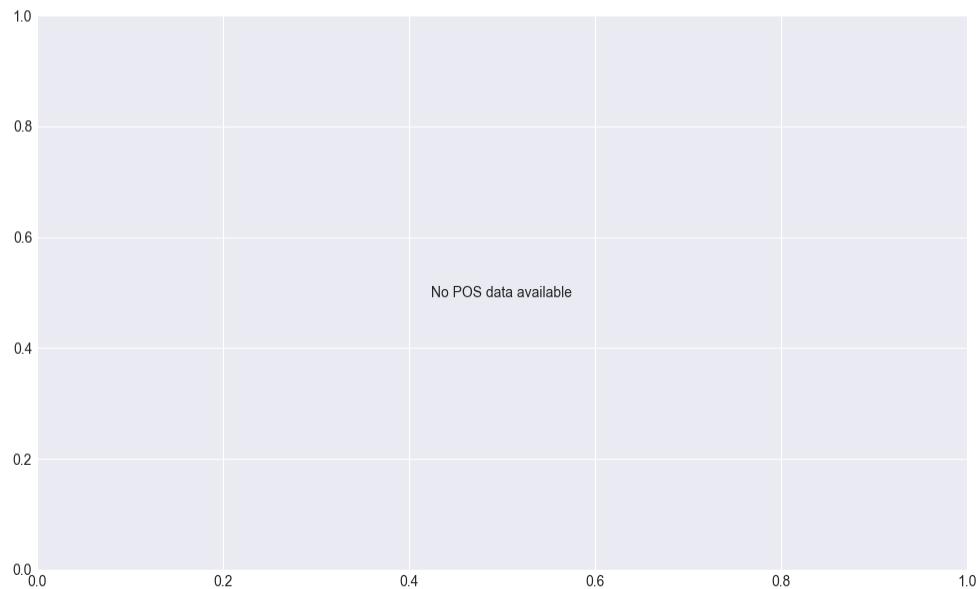
Sentence Length Distribution



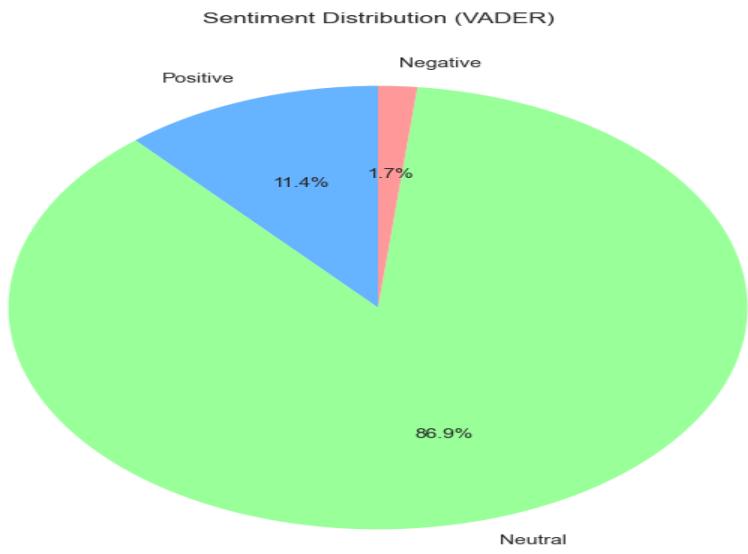
Zipf's Law Analysis (Word Frequency vs Rank)



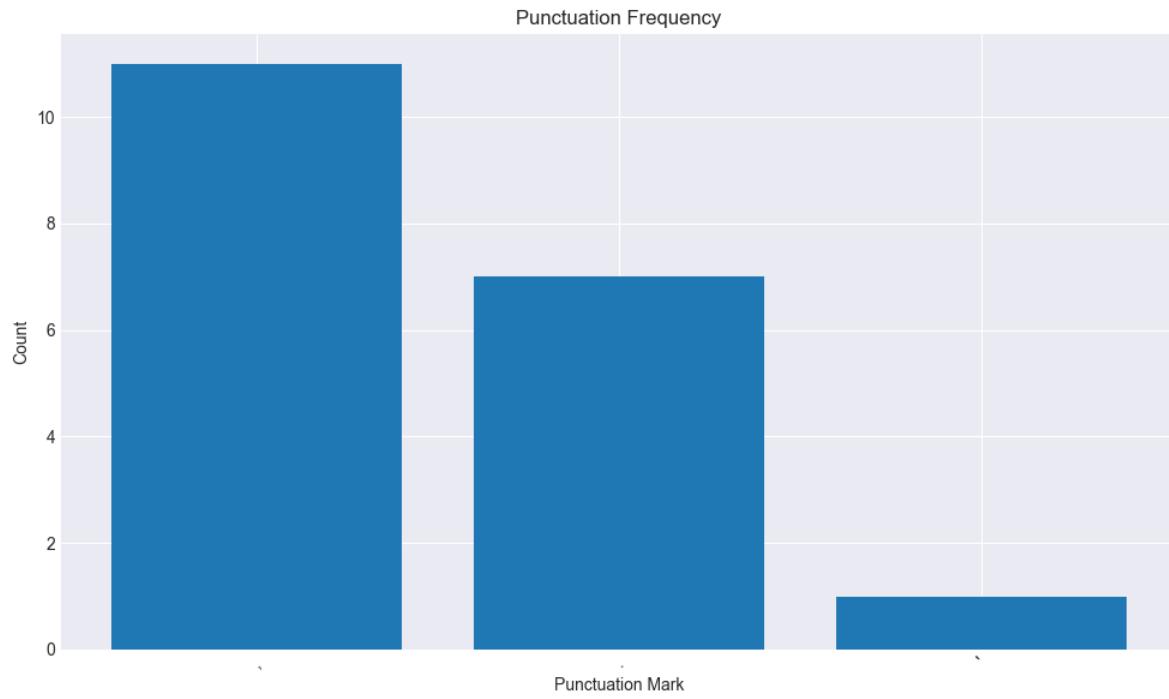
Part-of-Speech Distribution (Top 25)



Sentiment Distribution (VADER)



Punctuation Frequency



5. Frequency Analysis

Most Common Words (Top 15, with stopwords):

blockchain (6), the (5), technology (4), a (4), transactions (4), in (4), and (4), as (3), bitcoin (3), of (3), is (2), decentralized (2), ledger (2), that (2), such (2)

Most Common Punctuation:

',' (11), '.' (7), '"' (1)

6. N-gram Analysis

Common 2-grams (Top 10):

'blockchain technology' (2), 'blockchain is' (2), 'a decentralized' (2), 'of blockchain' (2), 'technology blockchain' (1), 'is a' (1), 'decentralized distributed' (1), 'distributed ledger' (1), 'ledger technology' (1), 'technology that' (1)
2-gram Repetition Rate: 0.0310

Common 3-grams (Top 10):

'blockchain technology blockchain' (1), 'technology blockchain is' (1), 'blockchain is a' (1), 'is a decentralized' (1), 'a decentralized distributed' (1), 'decentralized distributed ledger' (1), 'distributed ledger technology' (1), 'ledger technology that' (1), 'technology that records' (1), 'that records transactions' (1)
3-gram Repetition Rate: 0.0000

Common 4-grams (Top 10):

'blockchain technology blockchain is' (1), 'technology blockchain is a' (1), 'blockchain is a decentralized' (1), 'is a decentralized distributed' (1), 'a decentralized distributed ledger' (1), 'decentralized distributed ledger technology' (1), 'distributed ledger technology that' (1), 'ledger technology that records' (1), 'technology that records transactions' (1), 'that records transactions across' (1)
4-gram Repetition Rate: 0.0000

Character N-gram Analysis:

Common Character 3-grams (Top 10):

'in_-' (14), 'ed_-' (10), 'cha' (9), 'hai' (8), 'ain' (8), 'ion' (8), '_in' (8), 'blo' (7), 'loc' (7), 'ock' (7)

Common Character 4-grams (Top 10):

'chai' (8), 'hain' (8), 'ain_-' (8), 'bloc' (7), 'lock' (7), 'tion' (7), 'ockc' (6), 'ckch' (6), 'kcha' (6), 'acti' (6)

Common Character 5-grams (Top 10):

'chain' (8), 'hain_-' (8), 'block' (7), 'lockc' (6), 'ockch' (6), 'ckcha' (6), 'kchai' (6), '_tran' (5), 'trans' (5), 'tions' (5)