🧠 GraphJudge - Intelligent Knowledge Graph Construction System

GraphJudge is an intelligent knowledge graph construction system based on large language models. Through a three-stage processing pipeline, it extracts entities from Chinese text, generates knowledge triples, and uses AI for quality assessment.

💡 **Getting Started:** Upload a Chinese text file (.txt) or paste text directly to begin analysis.

Total Runs	Average Processing Time	Total Generated Triples	Recent Approval Rate
1	1242.6s	47	78.7%

📝 Input Text

Please upload a Chinese text file (.txt) or enter text directly:

📁 Upload File ✏️ Type Text

Drag and drop file here
Limit 200MB per file • TXT

Browse files

chapter1_raw.txt 23.4KB

✅ File uploaded successfully! (Encoding: utf-8)

File Name	File Size	Encoding Used	Chinese Chars
chapter1_raw.txt	23,996 bytes	utf-8	6,770

📋 File Content Preview

📊 Text Statistics

Character Count	Line Count	Paragraph Count	Est. Reading Time
8120	90	1	27.1 min

🚀 Start Processing 📊 View Detailed Results 📈 Historical Comparison

---

🏆 Final Results

✅ Approved Triples	❌ Rejected Triples	🎯 Average Confidence	⏱️ Total Processing Time
37	10	0.779	1242.6s
↑ 78.7% approval rate	↑ 21.3% rejection rate		

🧠 Final Knowledge Graph

After AI judgment, the following **37** knowledge triples were deemed accurate:

📋 Knowledge Triple Details

#	Knowledge Triple	AI Confidence	Quality Grade
1	【女媧氏】→地點→【大荒山】	0.800	🥉 Good
2	【女媧氏】→地點→【無稽崖】	0.800	🥉 Good
3	【石頭】→地點→【青埂峰】	0.800	🥉 Good
4	【石頭】→化為→【通靈寶玉】	0.800	🥉 Good
5	【女媧氏】→遺留→【石頭】	0.800	🥉 Good
6	【三生石】→地點→【西方靈河】	0.800	🥉 Good
7	【絳珠草】→地點→【三生石】	0.800	🥉 Good
8	【神瑛侍者】→灌溉→【絳珠草】	0.800	🥉 Good
9	【神瑛侍者】→地點→【赤瑕宮】	0.800	🥉 Good
10	【絳珠草】→化為→【絳珠仙子】	0.800	🥉 Good

🕸️ Relationship Network Graph

📊 Network graph requires Plotly library: `pip install plotly`

Text-based relationship display:

1. 女媧氏 → 地點 → 大荒山
2. 女媧氏 → 地點 → 無稽崖
3. 石頭 → 地點 → 青埂峰
4. 石頭 → 化為 → 通靈寶玉
5. 女媧氏 → 遺留 → 石頭
6. 三生石 → 地點 → 西方靈河
7. 絳珠草 → 地點 → 三生石
8. 神瑛侍者 → 灌溉 → 絳珠草
9. 神瑛侍者 → 地點 → 赤瑕宮
10. 絳珠草 → 化為 → 絳珠仙子
11. 三生石 → 因緣 → 絳珠草
12. 通靈寶玉 → 因緣 → 神瑛侍者
13. 通靈寶玉 → 因緣 → 絳珠仙子
14. 通靈寶玉 → 因緣 → 警幻仙子
15. 情僧 → 地點 → 青埂峰

📤 Export Options

📄 Export as JSON 📊 Export as CSV 📋 Generate Report

🔍 View Detailed Results by Stage

🔍 Entity Extraction Results

Status	Entities Found	Processing Time	Text Reduction
✅ Success	38	76.77s	85.1%
↑ Processing Complete	↑ 38 items	↑ GPT-5-mini API	↑ Denoising Effect

🔬 Detailed Processing Phases

Phase 1: Entity Extraction with GPT-5-mini

📝 **Advanced Language Understanding:** GPT-5-mini analyzes classical Chinese text using contextual understanding and entity recognition patterns optimized for Chinese literature.

📊 Extracted Entities with Smart Categorization

👤 **人物 (Characters)**

封氏 女媧氏

🏛️ **地點 (Locations)**

閶門 十里街 仁清巷 葫蘆廟

💭 **概念 (Concepts)**

紅樓夢 石頭記 通靈寶玉 石頭 甄士隱 賈雨村 英蓮 封肅

... and 24 more 概念

Phase 2: Text Denoising and Restructuring

🧹 **GPT-5-mini Text Optimization:** Intelligently restructures and cleans the text based on extracted entities, removing redundant descriptions while preserving essential factual content for accurate knowledge graph generation.

📄 **Original Input Text**

Raw Input

紅樓夢

第一回　甄士隱夢幻識通靈　賈雨村風塵懷閨秀

作者自云：因曾歷過一番夢幻之後，故將真事隱去，而借通靈之說，撰此《石頭記》一書也。故曰「甄士隱夢幻識通靈」。但書中所記何事，

Total length: 8,120 characters

✨ **Denoised & Structured Text**

Processed Output

三、石書入世、編述與名號流變
通靈寶玉之所歷，刻於石上；後有空空道人（後更號情僧）於青埂峰下得見石文，遂錄為《石頭記》。空空道人以情入色、見色生情，改題《情僧錄》。吳玉峰署題為《紅樓夢》，東魯孔梅溪題《風月寶鑑》。後曹雪芹...

Processed length: 1,209 characters

📈 Denoising Statistics

Compression Ratio	Entity Density	Processing Quality
0.15	2.375	High
↑ 85.1% reduction	↑ entities per word	↑ Optimized for KG

📊 Detailed Processing Metrics

🔧 **API Performance Analysis**


```
{
  "model" : "GPT-5-mini"
  "processing_time_ms" : "76774.4"
  "entities_extracted" : 38
  "success_rate" : "100%"
  "api_calls" : "2 (extraction + denoising)"
  "language_support" : "Advanced Chinese"
}
```


📏 **Quality & Content Metrics**


```
{
  "unique_entities" : 38
  "total_entities" : 38
  "deduplication_rate" : "0.0%"
  "text_structure_improved" : "Yes"
  "classical_chinese_optimized" : "Yes"
}
```


⏰ **Processing Timeline**

Phase	Duration	Status
Input Validation	0.1s	✅
Entity Extraction	38.4s	✅
Text Denoising	38.4s	✅
Quality Validation	0.1s	✅

🔗 Knowledge Triple Generation Results

Status	Processing Time	Triple Count	Avg Confidence	Unique Entities
✅ Success	536.85s	47	N/A	38
↑ Generation Complete	↑ GPT-5-mini API	↑ Knowledge Relations	↑ Quality Score	↑ Graph Nodes

Application error occurred

StreamlitAPIException: Expanders may not be nested inside other expanders.

Traceback:

```
File "D:\AboutCoding\AI_Research\GraphJudge_TextToKG_CLI\streamlit_pipeline\app.py", line 171, in run
    self._render_main_interface()
File "D:\AboutCoding\AI_Research\GraphJudge_TextToKG_CLI\streamlit_pipeline\app.py", line 349, in _render_main_interface
    self._render_results_section(st.session_state.current_result)
File "D:\AboutCoding\AI_Research\GraphJudge_TextToKG_CLI\streamlit_pipeline\app.py", line 551, in _render_results_section
    display_triple_results(result.triple_result)
File "D:\AboutCoding\AI_Research\GraphJudge_TextToKG_CLI\streamlit_pipeline\ui\components.py", line 398, in display_triple_result
    with st.expander("🔬 Detailed Triple Generation Phases", expanded=True):
         ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\USER\AppData\Local\Programs\Python\Python312\Lib\site-packages\streamlit\runtime\metrics_util.py", line 410, in w
    result = non_optional_func(*args, **kwargs)
             ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\USER\AppData\Local\Programs\Python\Python312\Lib\site-packages\streamlit\elements\layouts.py", line 601, in expan
    return self.dg._block(block_proto=block_proto)
           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "C:\Users\USER\AppData\Local\Programs\Python\Python312\Lib\site-packages\streamlit\delta_generator.py", line 518, in _block
    _check_nested_element_violation(self, block_type, ancestor_block_types)
File "C:\Users\USER\AppData\Local\Programs\Python\Python312\Lib\site-packages\streamlit\delta_generator.py", line 598, in _check_
    raise StreamlitAPIException(
```

Ask Google Ask ChatGPT