

Grammar Rules Rule

DocScanner AI Writing Rule Documentation

Rule Overview

Name: Grammar Rules

File: grammar_rules.py

Description: Checks for common grammatical errors and inconsistencies

Purpose: Ensure grammatical correctness and professional quality

Functions

`_get_nlp()`

`check(content)`

Regular Expression Patterns

```
) text_content = soup.get_text() doc = nlp(text_content) #
----- # Regex-based checks # -----
# Multiple consecutive spaces rule removed per user request # Double punctuation
check removed per user request # if re.search(r
^\s*(info|warning|note|tip|caution)\s*
^[a-z][a-z0-9_]*\s*:\s*
^[a-z][a-z0-9]*(/[a-z0-9_\-\.\ ]*)+/?$
^[a-z][a-z0-9_\-]*\.[a-z0-9]+$
```

Code Examples

```
■■ RAG rule helper not available - using standard grammar rules only
RAG helper not available for {__name__} - using basic rules
■■ spaCy model en_core_web_sm not found - grammar rules disabled
```

Dependencies

- logging
- title_utils.is_title_or_heading
- re
- spacy
- rag_rule_helper.check_with_rag
- bs4.BeautifulSoup

Implementation Notes

- Import RAG system with fallback

- Lazy load spaCy model to avoid Flask startup conflicts
- Increase max_length to handle large documents
- Lazy load spaCy model
- Extract clean text (strip HTML if present)

Generated on 2025-10-17 18:29:07
DocScanner AI - Writing Quality Enhancement System