# Terminology Rules Rule

*DocScanner AI Writing Rule Documentation*

## Rule Overview

**Name:** Terminology Rules
**File:** terminology_rules.py
**Description:** Validates technical terminology and preferred word usage
**Purpose:** Maintain consistent technical vocabulary and terminology

## Functions

**_get_nlp()**

**check(content)**

## Checks Performed

• Consistency Checking

## Regular Expression Patterns

```
) text_content = soup.get_text() try: nlp_model = _get_nlp() doc =
nlp_model(text_content) except Exception as e: # Fall back to basic processing if
spaCy fails doc = None # ---------------------------- # Regex-based terminology
checks # ---------------------------- for wrong, preferred in TERMINOLOGY.items():
pattern = re.compile(rf
```

## Code Examples

```
RAG helper not available for {__name__} - using basic rules
log in (verb) / login (noun)
set up (verb) / setup (noun)
```

## 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
• Load spaCy model lazily to avoid startup issues
• Increase max length to 3MB
• Example terminology dictionary (customize for your manuals)
• Extract plain text (strip HTML if present)