

Definition

“Slop” refers to AI-generated text that is low-quality. It can appear superficially correct but is some combination of generic, overly verbose, inaccurate, irrelevant to its intended purpose, and contributing little meaningful value to the reader (despite sounding fluent). “Slop” typically displays patterns of repetition, formulaic structure, vague language, and an absence of authentic perspective. Not all AI-generated text is “sloppy,” and human writing can be sloppy too.

Below, we outline all the key text characteristics (as codes) that contribute to “slop”. Please refer to this guide when labeling spans of text for “slop.” A span may fall under multiple codes.

Annotation Instructions

1. Read each text in full before making any annotations.
2. Make an initial assessment: Is this text “slop”? (yes/no)
3. If yes, highlight specific spans of text and assign appropriate code(s).

Please note that recognizing text as AI-written doesn’t automatically make it sloppy! You may find “slop” in parts of a text while still answering “no” to the overall “slop” assessment if these issues are minor.

General Guidance

- Focus on quality issues, not just AI detection.
- Consider the context, purpose, and intended audience of the text when making judgments.
- Code for the most significant issues rather than tagging everything possible.

- When in doubt about which code to use, choose the one that best captures the core problem.
- For ambiguous cases, briefly note your reasoning.

“Slop” Codes with Examples

Information Quality

IQ1: Factuality

- Incorrect or fabricated information
- Misleading or fallacious claims

Example: “Dr. Sarah Johnson of Harvard University published groundbreaking research on this topic in 2022.” (Slop if Dr. Johnson doesn’t exist, isn’t at Harvard, or didn’t publish such research)

IQ2: Bias

Lack of appropriate perspective or over-standardization.

- Text that feels too “objective” when subjectivity is appropriate
- Missing rhetorical point of view when one is needed
- Lack of appropriate perspective based on the context
- Content that seems detached when engagement is required
- The presence of an inappropriate perspective or inappropriate assumptions

Example: “The economic policy changes of 2023 were universally beneficial.” (Slop because it presents a one-sided view of complex policy impacts)

Information Utility

IU1: Information Density

- Text that is verbose but conveys little actual information
- Generic statements that could apply in almost any context
- Excessive filler words and phrases that add no value

Example: “In today’s fast-paced modern world of cutting-edge technology and innovation, it has become increasingly important to consider the various factors and elements that contribute to our understanding of this complex and multifaceted issue.” (Slop because it uses many words to say almost nothing)

IU2: Information Relevance

- Appropriateness to the specific context, query, or task.
 1. Content that fails to address the nuances of the query or task
 2. Content that contributes nothing meaningful to context/query/task
 3. Text that appears disconnected from its intended purpose
 4. For text with additional context (e.g., query/prompt), consider relevance to such texts
 5. For text with no additional context (e.g., an article), consider internal relevance within the passage
- Example: In response to “How can I improve my marathon time?”: “Running is an excellent form of exercise with many health benefits including improved cardiovascular function, enhanced mood, and weight management.” (Slop because it doesn’t address the specific question about improving marathon times)

Style Quality

SQ1: Repetition

1. Excessive use of the same words or phrases
2. Redundant statements that add no new information
3. Overuse of transitional phrases common in formulaic writing
4. Low diversity in vocabulary and expression

Example: “The project was a success. The team accomplished their goals successfully. The successful outcome was due to the team’s hard work.” (Slop due to repetition of “success/successful” without adding new information)

SQ2: Templatedness

- Over-reliance on formulaic structures and patterns.
 1. Predictable formatting patterns (e.g., excessive use of bullet points)
 2. Standard transitional phrases used repeatedly
 3. Frequent appearance of text that follows a common pattern (e.g., “Mr. X, a Y-year-old Z”)

Example: “Dr. Smith, a researcher at Oxford University, found that... Professor Johnson, a scientist at Cambridge University, discovered that... Dr. Williams, an expert at Yale University, confirmed that...” (Slop because it follows the same formula repeatedly)

SQ3: Coherence

- Logical flow and structural integrity of the text.
 1. Poor sentence structure or organization
 2. Inconsistencies in argument or narrative
 3. Text that requires significant effort to follow
 4. How paragraphs work together to advance the argument or story

Example: “Climate change is affecting global temperatures. Polar bears are mammals. Ice cream melts in warm weather. Arctic ice is melting. Some people enjoy winter sports.” (Slop because the sentences, while related to temperature, don’t flow logically)

SQ4: Language Naturalness

- Formerly “Fluency” - Language that sounds artificial or manufactured.
 1. Strange turns of phrases or unnatural language
 2. Technically correct grammar that still reads unnaturally
 3. Misaligned word choice for the context
 4. Can co-occur with verbosity if long, does not necessarily include complex words

Example: “The earthen area that formerly held the puddle was now dry.” (Slop because natural language would simply say “The puddle had dried up” or “The ground where the puddle had been was now dry”)

How Language Naturalness differs from related codes:

- (vs. Verbosity)
 - Language Naturalness focuses on awkward phrasing even in short sentences
 - Concerns how humans naturally express ideas, not just length
 - Can occur in concise text that still sounds machine-generated or translated
- (vs. Word Complexity)
 - Language Naturalness can involve simple words arranged unnaturally
 - Focuses on sentence structure and phrasing patterns
 - Concerns how words fit together, not just vocabulary choices

SQ5: Verbosity

- Excessive wordiness relative to the information conveyed.
 1. Unnecessarily “flowery” or descriptive language
 2. Text that prioritizes word count over meaningful content
 3. Long-winded explanations that need significant editing

Example: “The consumption of the aforementioned beverage, which had been prepared with the utmost care and attention to detail by the skilled barista, provided me with a sense of satisfaction and contentment that permeated my entire being.” (Slop because it could simply say “I enjoyed the coffee”)

SQ6: Word Complexity

- Inappropriate use of vocabulary relative to context.
 1. Unnecessary jargon or complicated terminology
 2. Content filled with buzzwords that obscure meaning
 3. Overuse of rare words

Example: In a general article about gardening: “The phenolic compounds in certain cultivars exhibit antimicrobial properties that mitigate pathogenic microorganism

colonization.” (Slop because it uses unnecessarily complex terminology for the intended audience)

SQ7: Tone

- Appropriate voice and style for the context.
 1. Generic voice lacking character or purpose
 2. Missing perspective or point of view
 3. Overly formal language in casual contexts (or vice versa)
 4. Text that reads like an outside observer rather than engaged writer
 5. Overconfidence in response (sometimes with IQ1)
 6. Can have a relationship with factuality (IQ1)

Example: In a blog post about personal travel experiences: “The aforementioned destination offers numerous recreational activities for tourists. Visitors may engage in swimming, hiking, or dining at local establishments.” Slop because it uses an inappropriately formal tone for a personal blog)