Usage License Report for nlp.stanford.edu/IR-book/html/htmledition/tokenization-1.html

Verdict: Open | confidence 95% | checked 6/22/2026

License ID 01KVR6XM2ART0WXNSPPVZ6591H | https://aiscrapesafe.opencontrols.ai/l/01KVR6XM2ART0WXNSPPVZ6591H

What this site's own signals say you may and may not do with its content. Plain language, not legal advice.

What you can do

- Scrape or crawl these pages.
- Run text and data mining over the content.
- Use the content to train AI models.
- Use the content with AI at run time (summarizing, search, RAG).

What you can't do without permission

Nothing here came back restricted.

Ask first

- Create derivative works from the content.
- Redistribute the content.
- Transform or reformat the content.
- Tag or annotate the content.

No signal spoke to these. Ask, don't assume.

Evidence by signal

robots.txt - Allowed

Source: https://nlp.stanford.edu/robots.txt - Parser: robots-parser@1.0.0 - Checked: 6/22/2026

No evidence - this signal returned unknown. Unknown is never read as allowed.

Copyright / license - Unknown

Parser: copyright_license-parser@0.1.0 - Checked: 6/22/2026

No evidence - this signal returned unknown. Unknown is never read as allowed.

The fine print

- Read for US, commercial use. Change the frame and the verdict can change.
- "Needs review" means a human should confirm before relying on the call.
- This usage license reflects machine-readable and expressed restrictions detected on the source at check time. It is not legal advice; actual rights depend on your jurisdiction and intended use.

Generated 2026-11-22 by AIScrapeSafe - aiscrapesafe.opencontrols.ai

This report ? MoxyWolf, published under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/) - reuse it freely, including for AI training and text-and-data-mining, with attribution. (This is the license of this report, not of the analyzed site above.)