Greenwashing the Future? Computational Text Analysis of Environmental Reporting from the Fossil Fuel Industry

Robin Rauner

This README file lists the R scripts used to produce the tables and figures presented in the main paper and Online Supporting Information. All scripts were run successfully on 2 January 2026.

00_reproduce_analysis.R

Runs all scripts in order and creates HTML log files. Ensure that all required packages are installed prior to running the script.

01_dictionary_analysis.R

Input:

- `messages_text_corpus.csv`
- `fls_terms_no_glob.csv`

Output:

- `dictionary_annotated_data.csv`

02_performance_metrics.R

Input:

- `validation_set.csv`
- `held_out_validation_set.csv`
- `validation_set_coder_2.csv`
- `held_out_validation_set_coder_2.csv`

- handcoded_set_R.csv
- handcoded_set_K.csv
- llm_annotated_data.csv
- dictionary_annotated_data.csv

Output:

- performance_metrics_test_heldout.tex: Table A4
- performance_metrics_n250.tex: Table 1

03_calculate_FLS_score.R

Input:

- llm_annotated_data.csv
- messages_text_corpus.csv

Output:

- fls_scores_llm.csv

04_analysis.R

Input:

- llm_annotated_data.csv
- fls_scores_llm.csv
- report_stats.csv
- nzt_data.xlsx: Net Zero Tracker data
- new_docvars.csv
- Company_Latest_Assessments_5.0.csv: Transition Pathway Initiative (TPI) sector data
- fls_terms_all.csv

Output:

- corpus_overview.tex: Table A1
- keyness.pdf: Figure 3
- frequency.pdf: Figure A1
- report_type_features.tex: Table A6
- fls_terms_chi2.tex: Table A5
- fls_scores_by_region.pdf: Figure 4
- fls_scores_mean_sd.pdf: Figure A2
- fw_matrix_plot.pdf: Figure 5
- regressions_future_focus.tex: Table A7