Vox Populi, Vox AI?
Using Language Models to Estimate German Public Opinion

Leah von der Heyde, Anna-Carolina Haensch, Alexander Wenz

Supplementary Material

Codebook for LLM data collection files

Orange: Output data from completions
Yellow: Input data from (=after) manual checks for resampling
Green: Clean final data

1. **Main Dataset**

- manualchecks: mc_GLES2017, merged GLES & first round of completions, unchecked → for manual checks
  - manual_checked: mc_GLES2017, merged GLES & first round of completions, checked, matches only → merge to final
  - manual_resample: mc_GLES2017, merged GLES & first round of completions, checked, NAs only → input for resampling
    - manualchecks_resample: match_completions_resample, unmerged, second round of completions, unchecked → for manual checks
      - manual_resample_checked: unmerged, second round of completions, checked → archived, only saved for replication (being able to re-run script)
      - manualchecks_resample_merged: mc_resample, merged resample & second round of completions, checked
      - manual_resample_merged_checked: mc_resample, merged GLES & second round of completions, checked, matches only → merge to final
      - manual_resample_2: mc_resample, merged GLES & second round of completions, checked, NAs only → input for second resampling
        - manualchecks_resample_2: mc_resample_2, merged GLES & third round of completions, unchecked → for manual checks
          - manual_resample_2_checked: mc_resample_2, merged GLES & third round of completions, checked → merge to final

2. Robustness: Non-imputed prompts

- manualchecks_robustness: mc_GLES2017_nonimputed, merged GLES & first round of non-imputed completions, unchecked → for manual checks
  - manual_robustness_checked: mc_GLES2017_nonimputed, merged GLES & first round of non-imputed completions, checked, matches only → merge to final
  - manual_robustness_resample: mc_GLES2017_nonimputed, merged GLES & first round of non-imputed completions, checked, NAs only → input for resampling
    - manualchecks_resample_robustness: mc_resample_nonimputed, merged resample & second round of non-imputed completions, unchecked → for manual checks
      - manual_resample_checked_robustness: mc_resample_nonimputed, merged GLES & second round of non-imputed completions, checked, matches only → merge to final
      - manual_resample_2_robustness: mc_resample_nonimputed, merged GLES & second round of non-imputed completions, checked, NAs only → input for second resampling
        - manualchecks_resample_2_robustness: mc_resample_2_nonimputed, merged GLES & third round of non-imputed completions, unchecked → for manual checks
          - manual_resample_2_checked_robustness: mc_resample_2_nonimputed, merged GLES & third round of completions, checked → merge to final