

Safe AI Use for Qualitative Analysis Checklist

This infographic outlines the three zones for using AI responsibly when analyzing employee feedback. The zones indicate which AI tasks are permitted, restricted, or prohibited. Begin by checking where your task falls, then review the listed tasks and conditions to ensure responsible, ethical, and compliant use.

CHECK

Is it descriptive and at group-level?

GREEN ZONE

AI may be used to clean, organize, and summarize anonymized or deidentified qualitative data without interpreting intent, causes, or individual meaning.

Tasks

- Text cleaning and standardization
- Removing duplicates and repeated comments
- Clustering similar phrases or topics
- Producing descriptive summaries of what is mentioned most often
- Offering a high-level, descriptive sense of tone across a dataset, without attributing meaning or impact
- Drafting neutral theme labels or codebook definitions
- Rewriting findings for communication or reporting purposes, without adding interpretation

Conditions

- Report aggregated findings (avoid small groups; aim for 5–10+ per group).
- Keep an audit trail of what AI did and what humans changed.
- Never use green outputs for individual performance management.

CHECK

Is it interpretive, predictive, or uses personal data?

AMBER ZONE

AI may support exploratory analysis and pattern detection with safeguards. All outputs are provisional and require human judgment and validation.

Tasks

- Thematic coding and automated tagging at scale (humans decide final codes)
- Sentiment analysis by topic to spot trends (humans validate against raw text)
- Highlighting potential outliers/minority voices for human follow-up
- Scenario ideas for strategy discussions (treated as thought starters)
- Population-level predictive signals (team/department trends only, not individuals)

Conditions

- Requires human-in-the-loop review.
- Do not take direct action on individuals based solely on AI outputs.
- Conduct regular bias and error audits.
- Tell employees AI is used; how and what results are used for.

CHECK

Could it judge, score, manipulate, or decide about people?

RED ZONE

AI must not be used for tasks that judge, predict, or influence outcomes for individuals, even when data is anonymized or aggregated.

Tasks

- Inferring emotional or psychological states from employee feedback
- Inferring protected or sensitive personal characteristics from language or behavior
- Creating employee scores, rankings, or profiles that affect treatment or decisions
- Making significant employment decisions without meaningful human oversight
- Combining feedback data across contexts to profile employees
- Manipulating how employees respond to surveys or feedback mechanisms

Conditions

If an AI output could influence how a person is evaluated, treated, or managed, it does not belong in the green or amber zone and therefore is a red zone area.