

GENAI-POWERED ANALYSIS OF FREE-RESPONSE SURVEYS



TARLETON STATE
UNIVERSITY®

Member of The Texas A&M University System

Turning open-ended feedback into actionable
insight, overnight.

TAYLOR HUTYRA

#BLEEDPURPLE



* Please rate your agreement with each of the following statements:

	Strongly disagree	Somewhat disagree	Neutral	Somewhat agree	Strongly agree
Question 45	<input type="radio"/>				
Question 46	<input type="radio"/>				
Question 47	<input type="radio"/>				
Question 48	<input type="radio"/>				
Question 49	<input type="radio"/>				

* Please rate your agreement with each of the following statements:

	Strongly disagree	Somewhat disagree	Neutral	Somewhat agree	Strongly agree
Really? Another Survey	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
How long have I been doing this?	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
What kind of questions are these?	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Will my thoughts be answered anywhere??	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Brain. Tired. Must. Coffee	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Next 

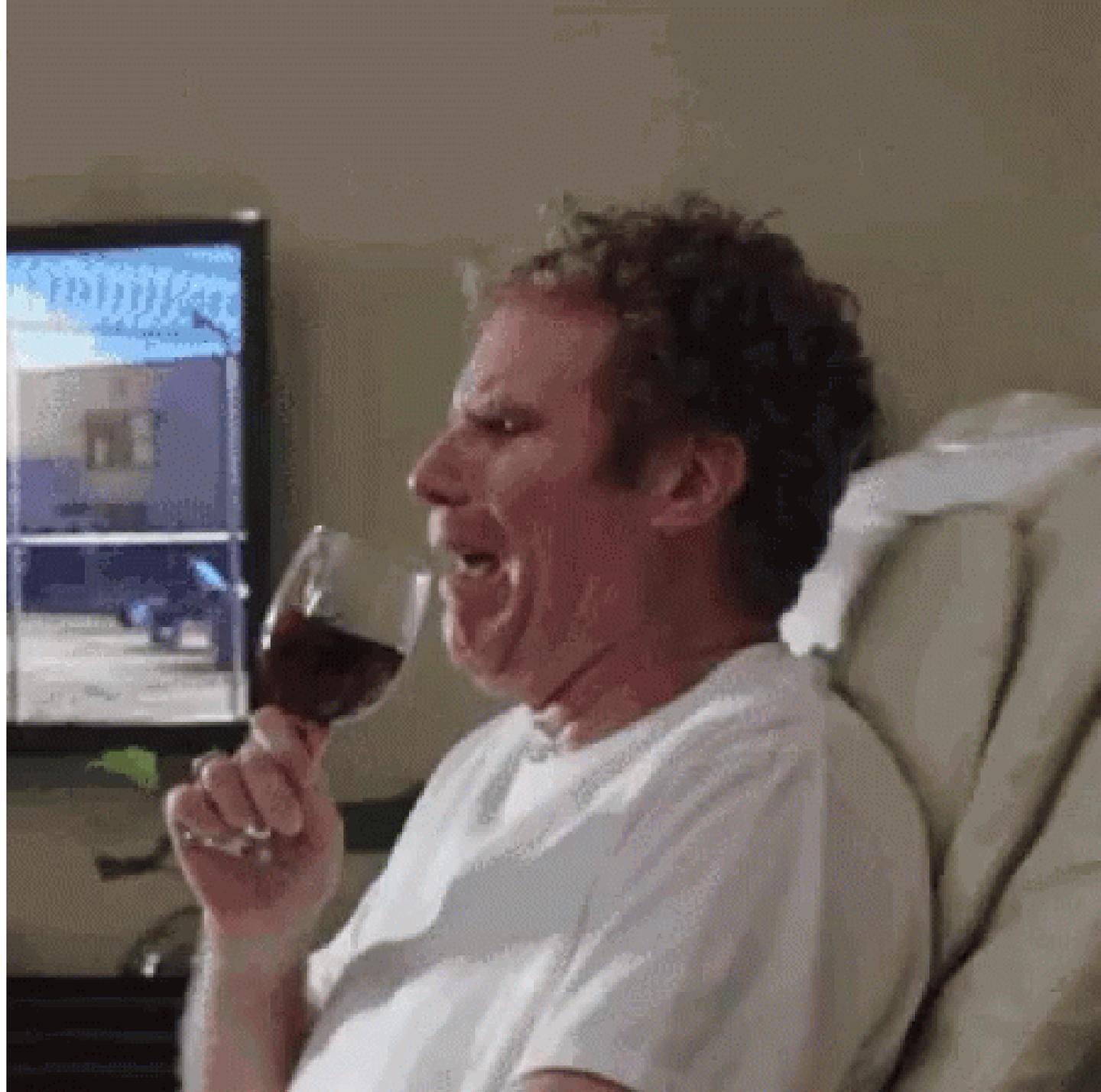
48%

Time Spent: 17:23

* Please rate your agreement with each of the following statements:

	Strongly disagree	Somewhat disagree	Neutral	Somewhat agree	Strongly agree
Really? Another Survey	<input type="radio"/>				
How long have I been doing this?	<input type="radio"/>				
What kind of questions are these?	<input type="radio"/>				
Will my thoughts be answered anywhere??	<input type="radio"/>				
Brain. Tired. Must. Coffee	<input type="radio"/>				

Next >



TYPICAL APPROACH : GUIDED SURVEYS

- **Likert scales** : 1–5 , Strongly Disagree → Strongly Agree
- **Multiple choice** : select all that apply, ranking
- These are the default. Easy to tabulate, easy to chart
- This is what most of us would reach for first



PROBLEMS: RESPONSE BIAS

- You already know the questions. You're **confirming hypotheses**, not discovering [1]
- **Acquiescence bias** —people lean toward “agree” out of politeness or fatigue [2]
- **Central tendency bias** —people avoid extremes, cluster in the middle [3][4]
- **Social desirability bias** —answers shaped by what feels “acceptable” [5]
- **Narrow scope** —if the real issue isn't in your predefined options, you'll never find it [1]



PROBLEMS: SURVEY FATIGUE

9%

take their time

67%

have abandoned a
survey midway

8 min

max before significant
drop-off begins

A survey is only as good as the data it collects. If your participants are checked out; the information is useless!



PROBLEMS: ANALYSIS & INSIGHTS

Likert can tell you what is happening but never tells you why.

3.2

Communication Satisfaction



OUR SCENARIO

- New interim provost after a sudden leadership transition
- Wants to hear from all faculty and staff:
 - what's working?
 - what's not?
 - what needs attention?
- Ground rules: keep it big-picture. Focus on issues and ideas, not specific people, positions, departments, or offices
- So, we need to do a survey. What kind?



OUR APPROACH: OPEN-ENDED QUESTIONS

4 simple questions:

Q1

What's working well?

Q2

What is not working well?

Q3

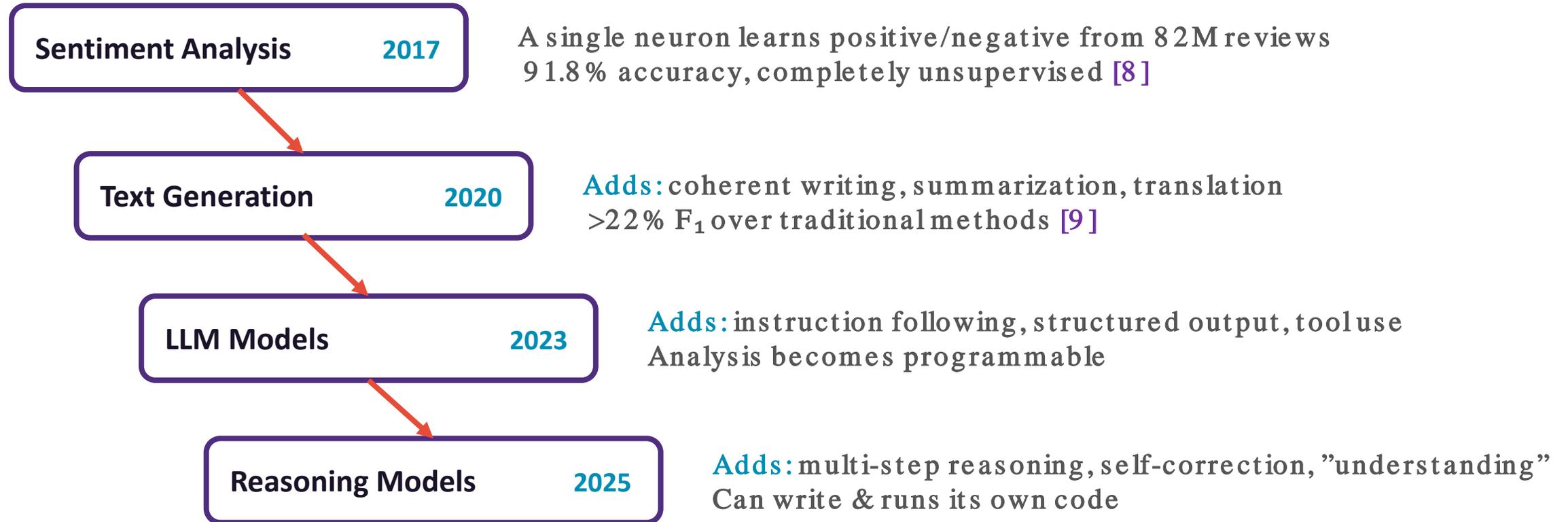
What are your hopes for our future?

Q4

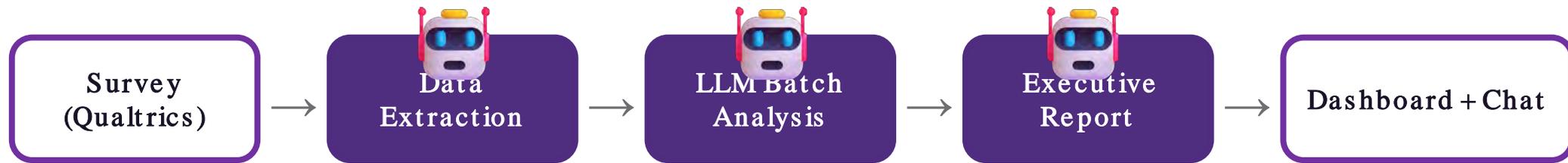
Anything else you want me to know?



FROM SENTIMENT NEURON TO STRUCTURED ANALYSIS



ARCHITECTURE OVERVIEW



Processing

Python and OpenAI API with structured outputs (GPT-5)

Visualizations

Dash + Plotly provide 8 interactive chart types

Chat Interface

Streamlit for prototyping natural language Q&A over survey data



DEMO



WHAT THE LLM EXTRACTS PER RESPONSE

Every individual response is analyzed in isolation, producing structured data across multiple dimensions.

Topics

Multiple topic labels per response: "faculty morale," "workload," "compensation"

Sentiment

Numeric score (-1 to +1) plus descriptive emotion labels per topic

Entities

Programs, committees, initiatives, and offices mentioned in context

Actionable Signals

Risks, opportunities, and urgency levels flagged for leadership

Key Quotes

Representative excerpts pulled verbatim for grounding and evidence

Structured JSON

All output validated against strict schemas for consistency across responses



REDACTION ENGINE

Our Provost's requirement: hear the concerns, protect the people.

“[NAME] in [DEPARTMENT] told me the [COMMITTEE] hasn't met in a while.”

- **Customizable** redaction categories, tuned per survey to balance anonymity and insight
- **Allowlist** for institutional terms that should not be redacted
- **Redaction at analysis time** —dashboard and chat only see redacted text



EXECUTIVE REPORT GENERATOR

- The LLM runs Python to directly query data.
- Tool calling enables real code execution on actual data.
- Python calculates all counts, percentages, and staff/faculty details.
- The LLM crafts narratives with verified stats and exact quotes.
- Outcome: a reliable executive summary with accurate numbers and authentic quotes.



FROM THE PROVOST

“I wanted a free response survey so I could hear what faculty really think is important. This tool let me dig into their responses much deeper than I ever could before and find subtle patterns I would never otherwise see. It let me engage directly with their concerns and immediately establish open and honest lines of communication.”



QUESTIONS?

Scan the QR code to
complete the session
survey.



2026



THANK YOU!



REFERENCES

- [1] Nielsen Norman Group. "Open-Ended vs. Closed Questions in User Research."
- [2] SmartSurvey. "Response Bias in Surveys: Types, Causes & How to Avoid It."
- [3] Joshi, A. et al. (2015). "Likert Scale: Explored and Explained." *British Journal of Applied Science & Technology*, 7(4), 396–403.
- [4] Sullivan, G. & Artino, A. (2013). "Analyzing and Interpreting Data From Likert-Type Scales." *Journal of Graduate Medical Education*, 5(4), 541–542.
- [5] Bergen, N. & Labonté, R. (2020). "Everything Is Perfect, and We Have No Problems." *Qualitative Health Research*, 30(5), 783–792.
- [6] Zoho Survey. "Survey Fatigue: What It Is and How to Prevent It."
- [7] Karlan, D. et al. (2022). "Exhaustive or exhausting? Evidence on respondent fatigue in long surveys." *Journal of Development Economics*, 161, 102992.
- [8] Radford, A. et al. (2017). "Learning to Generate Reviews and Discovering Sentiment." arXiv:1704.01444.
- [9] Liao, W. et al. (2024). "Optimization Techniques for Sentiment Analysis Based on LLM (GPT-3)." arXiv:2405.09770.

