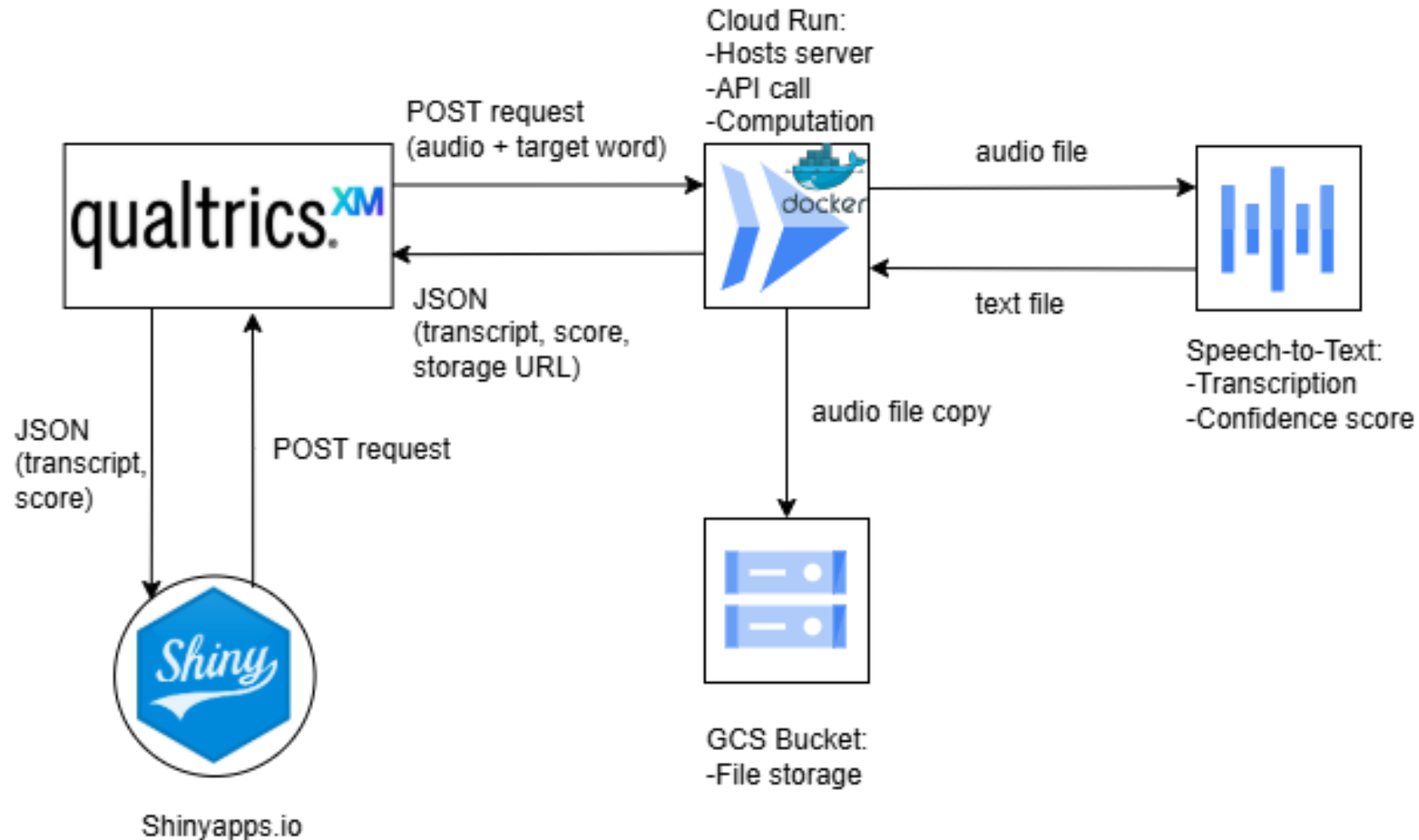# Realtime Audio-based Survey Response Processing Project

Jay Kim

Back End: Data Pipeline

Proximity Score Calculation

- Levenshtein Distance (insertion, deletion)
- (1 - levDistance / maxLen) * 100

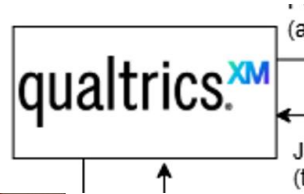   "i said apple" vs "apple"

    1 - 0.444 = 0.556 (55.6% similar)

    "cat" vs "dog"

    1 - 1.0 = 0 (0% similar)

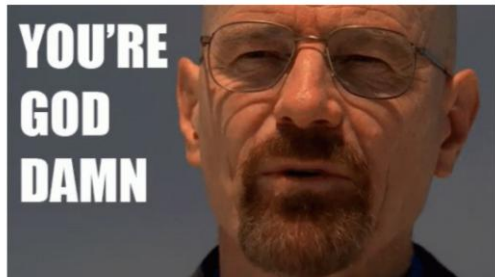|		k	i	t	t	e	n
	0	1	2	3	4	5	6
s	1	1	2	3	4	5	6
i	2	2	1	2	3	4	5
t	3	3	2	1	2	3	4
t	4	4	3	2	1	2	3
i	5	5	4	3	2	2	3
n	6	6	5	4	3	3	2
g	7	7	6	5	4	4	3

Front End

Press the record button and say his name.

qualtrics.XM

I think you said: *You haven't said anything yet*

▶ Start Recording   ■ Stop Recording

Ready to record

PS > 0.9

Bingo!

YOU'RE GOD DAMN

PS ≤ 0.9

Whaaat??

Survey Response Dashboard

Shiny

Controls

Select Survey:

SM727 BB survey   ▼

Refresh Data

Total responses: 8

Word Cloud Options

Max Words:

10   100   200

10 29 48 67 86 105 124 143 162 181 200

Min Frequency:

1   10

1 2 3 4 5 6 7 8 9 10

Transcript Word Cloud

eisenberg
apple
heisenberg
mber

ProximityScore Distribution

ProximityScore Distribution

Mean: 34.83

Median: 13.79

Frequency / Score