

UNITED

# Tackling Customer Reviews with Agentic AI

*Team 6 - The Infinite Loops*



# Team Intro



**Toby Chiu**  
**Project Manager**

---

UChicago MS-ADS Alum



**Chris Tomaszewicz**  
**Data Scientist**

---

UChicago MS-ADS Alum  
Customer Insights @ United  
Airlines



**Prachi Gyanmote**  
**Data Scientist**

---

Digital Tech Intern @ United  
Airlines



**Reethesh Venkataraman**  
**Data Scientist**

---

UChicago MS-ADS 25'



**Liang Gong**  
**Data Scientist**

---

UChicago MS-ADS 25'



**Rithik Alluri**  
**Data Scientist**

---

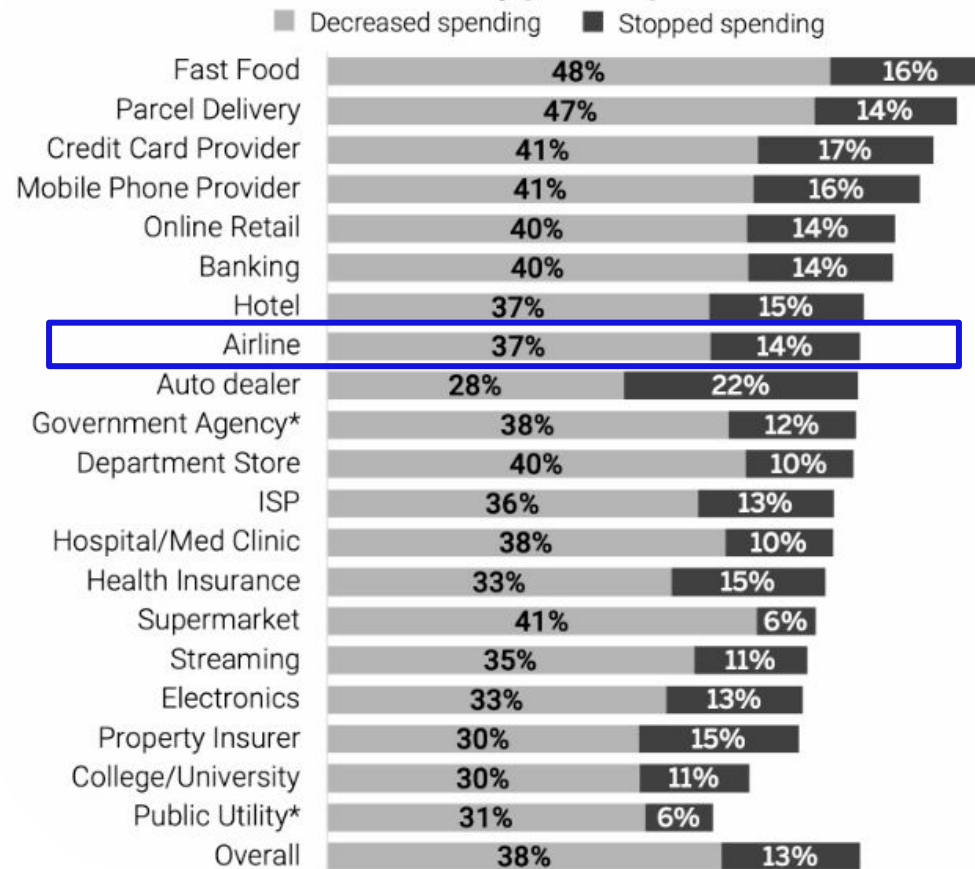
UChicago MS-ADS Alum



# Problem Statement

## Negative Consumer Experiences Can Lead to a Loss of Customers

Percentage of consumers who decreased or stopped spending after a 'very poor' experience



*Airline consumers with very poor experiences **decrease or stop spending** with the organization **51% of the time***

*70% of United's customers fly **once per year** - first impressions matter most*

*Customers leave after a bad experience, leading to **increased churn, lost market share, and lower revenue***



# Business Case

Problem

Solution

Considerations

Demo!

## Empower United to Address Customer Experience Issues


**Solution:** An **agentic chatbot tool** that lets them...


- **Thematic Trends:** Identify trends in reviews (both praise and complaints) over time
- **Deep Dive:** Dig deeper into specific review types and isolate what customers are describing as areas for improvement
- **Identify at-risk customers:** Find which repeat customers have left poor reviews

**Goal:** Respond dynamically to review data

- Which future customers can we target a better experience towards for future retention?
- Which comment themes are most prevalent for the last month?
- What stations had the top 5 review scores the previous month?



 **User**  
For flights departing MSP on date June 24, please return customers who had a previous rating of less than 5 but greater than 1

 **AI** Tool Calling Agent

✓ Finished

✓ 5.5s

For flights departing MSP on June 24, the following customers had a previous rating greater than 1 but less than 5:

- Corina Del Valle (previous rating: 2)
- K Dixon (previous rating: 2)
- BJ Carter (previous rating: 4)
- J Trusky (previous rating: 3)
- Tumey Charles (previous rating: 3)

Tool used: future\_flights



# Our Solution

Problem

Solution

Considerations

Demo!

## 1 Data Pipeline

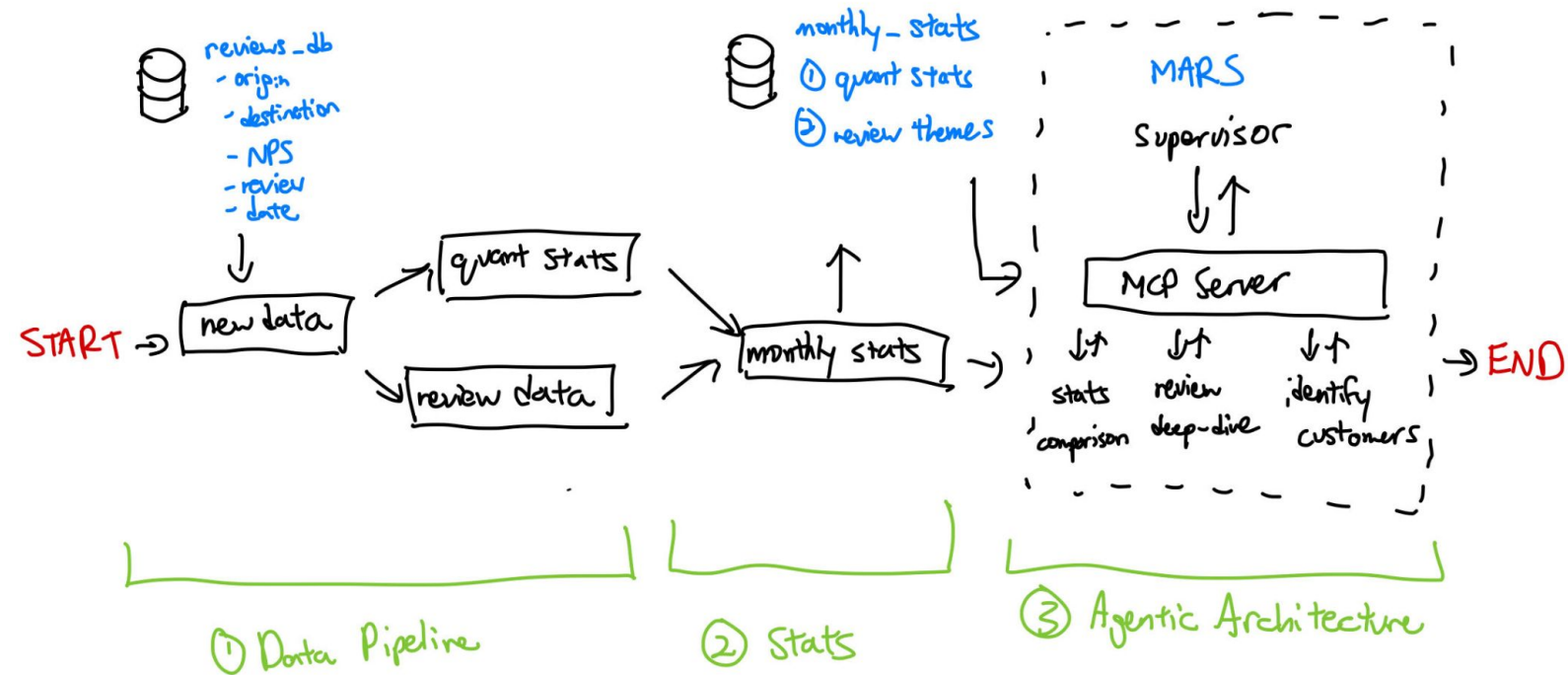
*Data Ingestion, Summary Statistics, Taxonomification*

## 2 Statistical Summary

*Database Design to allow for agentic interaction*

## 3 Agentic Architecture

*Supervisor <> MCP Server for toolkit*



# Data Ingestion Pipeline



Dataset

Filtering

LLM Imputation

Quantitative Scoring

Customer review data sourced  
from the publicly available  
**Skytrax Reviews Dataset.**

Includes airline names, textual  
feedback, and user-provided  
ratings across various flight  
experience dimensions.



# Data Ingestion Pipeline

Problem

Solution

Considerations

Demo!

Dataset

Filtering

LLM Imputation

Quantitative Scoring

Initial filtering **removes incomplete or irrelevant entries**, such as missing review content or invalid ratings.

Additional logic filters reviews to retain only those where either the **origin or destination is in the U.S.**



# Data Ingestion Pipeline

Problem

Solution

Considerations

Demo!

Dataset

Filtering

LLM Imputation

Quantitative Scoring

A Gemini-powered LLM **parses free-text reviews to extract structured origin and destination** IATA airport codes.

This step **enhances data quality by enriching missing route details** from unstructured content.





# Data Ingestion Pipeline

Problem

Solution

Considerations

Demo!

Dataset

Filtering

LLM Imputation

Quantitative Scoring

A **hierarchical taxonomy** of themes like timeliness, cleanliness, and staff behavior is applied to reviews.

Each review is scored based on theme presence and sentiment, then **aggregated by time**.



# Data - Quantitative

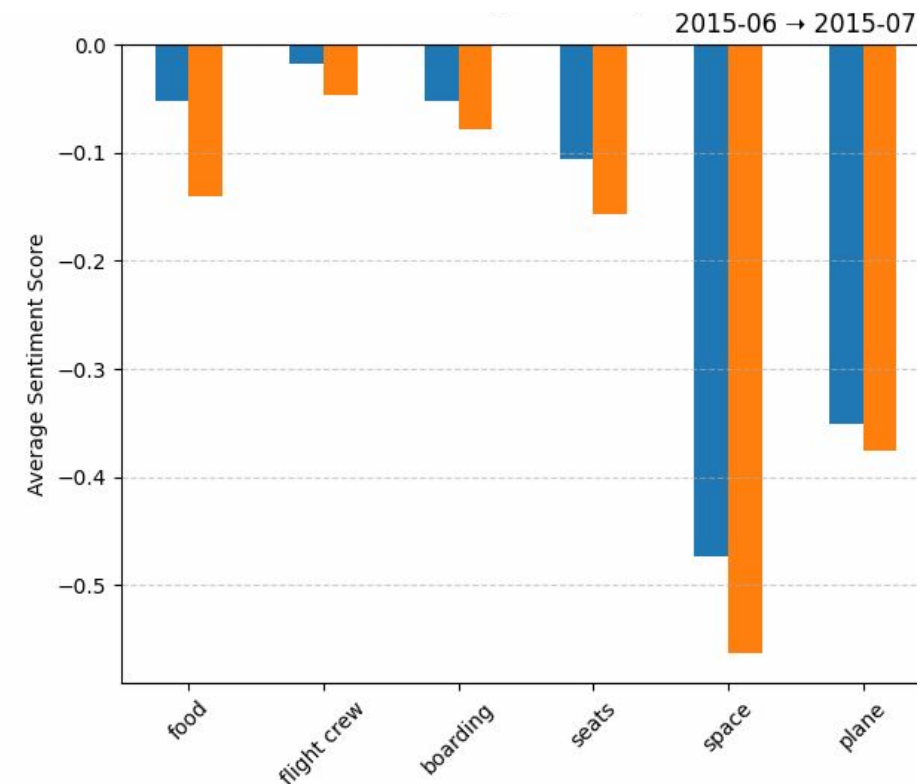
## Raw Data

Overall Rating	Seat Comfort Rating	Cabin Staff Rating	Food Beverages Rating	Inflight Entertainment Rating
7.0	4.0	4.0	4.0	0.0
10.0	4.0	5.0	4.0	1.0
9.0	5.0	5.0	4.0	0.0
8.0	4.0	4.0	3.0	1.0

## Quantitative Summary

```
{
  "last_month": "2015-07",
  "previous_month": "2015-06",
  "summary_last_month": [
    {
      "airline_name": "united-airlines",
      "overall_rating": 2.7,
      "seat_comfort_rating": 2.03,
      "cabin_staff_rating": 2.23,
      "food_beverages_rating": 1.8,
      "inflight_entertainment_rating": 1.9,
      "ground_service_rating": 2.03,
      "wifi_connectivity_rating": 1.93,
      "value_money_rating": 1.87,
      "recommended": 0.13,
      "number_of_reviews": 30
    }
  ],
  "month_to_month_comparison": [
    {
      "index": 90,
      "airline_name": "united-airlines",
      "cabin_staff_rating_delta": -0.24,
      "food_beverages_rating_delta": -0.04,
      "ground_service_rating_delta": -0.12,
      "inflight_entertainment_rating_delta": -0.26,
      "number_of_reviews_delta": 30.0,
      "overall_rating_delta": -0.67,
      "recommended_delta": -0.18,
      "seat_comfort_rating_delta": -0.39,
      "value_money_rating_delta": -0.24,
      "wifi_connectivity_rating_delta": -0.12
    }
  ]
}
```

## Category Rating Delta Month-Over-Month



# Data - Qualitative

Problem

Solution

Considerations

Demo!

## Raw Data

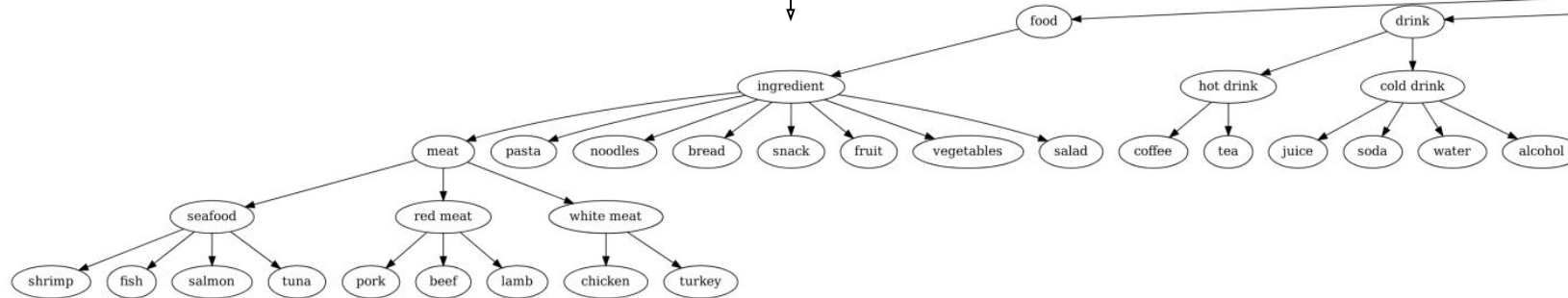
### Long-Form Review

Very fast on CRJ 90. Seats comfortable and crew was fine. Ground services good and lounge was quite nice.

Flew Zurich-Ljubljana on JP365 newish CRJ900. Flight almost full departure on time. Service on board by 2 pleasant and friendly female flight attendants. I bought coffee mineral water and a very nice and fresh filled croissant for EUR 6. Good value!

Adria serves this 100 min flight from Ljubljana to Amsterdam on a brand new CRJ900 Next Gen which has a hugely improved cabin.

## Taxonomy



## Sentiment Analysis

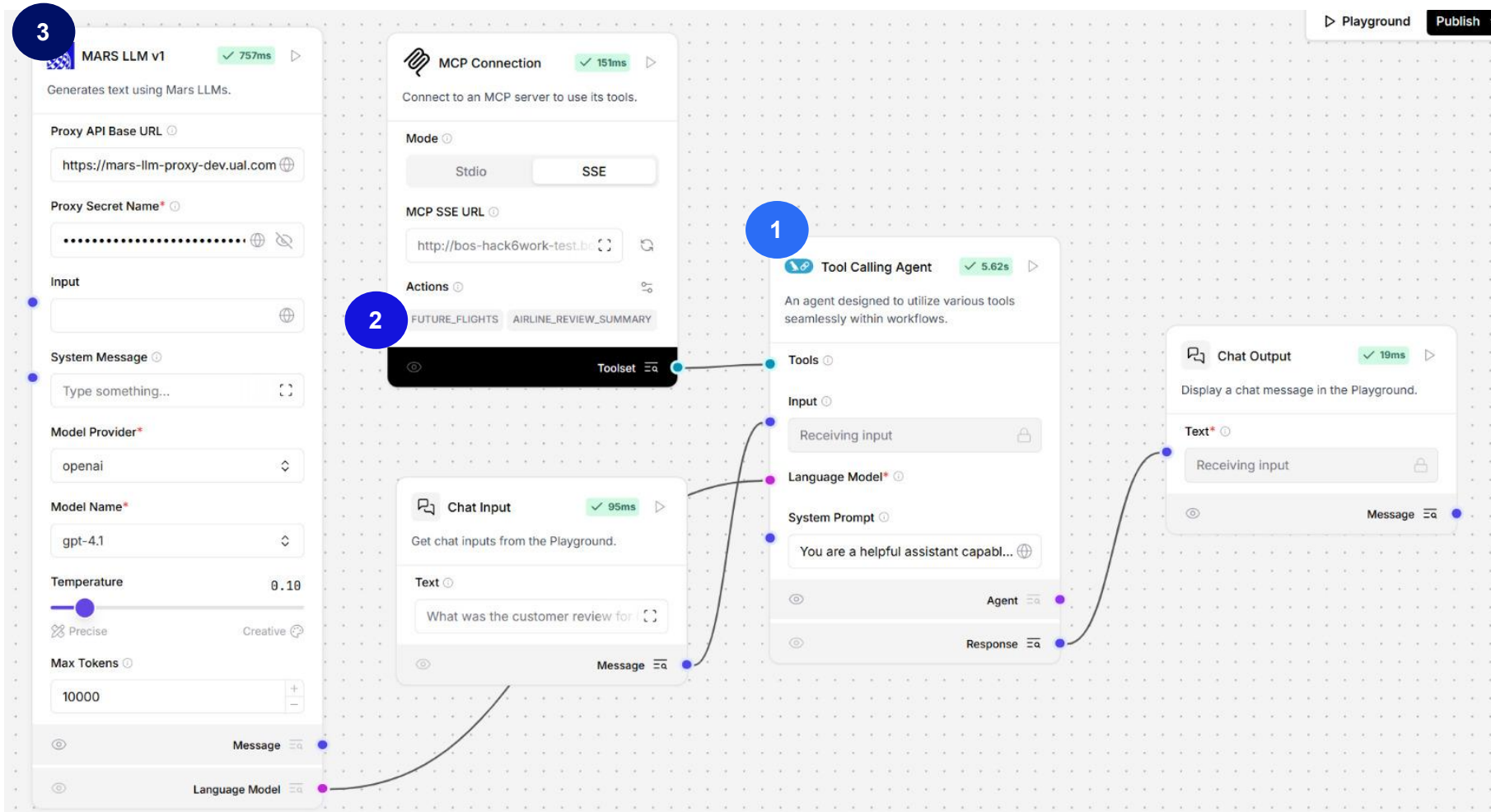
food	seats	service	plane	drink	flight crew	space	on time
0	1	1	0	0	0	0	0
1	0	1	1	1	1	0	1
0	0	0	1	0	0	1	0

## Qualitative Summary

```
{
  "last_month": "2015-07",
  "previous_month": "2015-06",
  "summary_last_month": [
    {
      "airline_name": "united-airlines",
      "ingredient": -0.02,
      "food": -0.14,
      "beef": 0.0,
      "meat": 0.0,
      "shrimp": 0.0,
      "drink": -0.17,
      "plane cleanliness": 0.0,
      "ljublana airport": 0.0,
      "snack": -0.05,
      "space": -0.56,
      "red meat": 0.0,
      "boarding": -0.08,
      "seats": -0.16,
      "fish": 0.0,
      "seafood": 0.0,
      "chicken": -0.02,
      "pasta": 0.0,
      "tea": -0.06,
      "flight": -0.94,
      "on time": -0.02,
      "coffee": -0.03,
      "pork": 0.0,
      "operations": -0.09,
      "service": -0.25,
      "flight crew": -0.05,
      "plane": -0.38,
      "number_of_reviews": 64
    }
  ],
}
```



# Agentic Architecture



## Commentary

## 1 Supervisor

- (a) *Receives user query and delegates to appropriate tool / action*
- (b) *Flags out-of-scope queries*

## 2 MCP Protocol

*Agent tools/actions:*

- (a) identifies at-risk customers*
- (b) time-period statistics*
- (c) thematic trends*

### 3 Chatbot

*Leverage OpenAI GPT-4.1 for generating response; keeps track of session history*



# Responsible AI

RAI Pillar	Potential Risk	Practical Mitigant
<b>Fairness</b>	<b>Representation skew</b> – highly vocal cities or demographics dominate summaries, hiding minority or low-volume voices	<ul style="list-style-type: none"> <li>• <b>Weighting</b> – Build a balanced evaluation set (region × demographic × sentiment)</li> </ul>
<b>Transparency / Explainability</b>	<b>Black-box insight</b> – summary numbers appear with no links to supporting review	<ul style="list-style-type: none"> <li>• <b>Customer Info</b> – Provide a “Why am I seeing this?” hover UI</li> </ul>
<b>Privacy &amp; Security</b>	<b>PII leakage</b> – raw reviews may contain passenger names, booking IDs, etc..	<ul style="list-style-type: none"> <li>• <b>PII Removal</b> – scrubber in ingestion pipeline (regex + ML redaction)</li> <li>• <b>Hashing</b> – Store hashed IDs; no raw emails in vector store</li> </ul>
<b>Safety &amp; Robustness</b>	<b>Out-of-scope answers</b> – user asks for about not related to customer reviews; agent complies	<ul style="list-style-type: none"> <li>• <b>Guardrails</b> – agent rejects non-review topics (“I’m only trained on customer-review insights”)</li> </ul>
<b>Accountability &amp; Governance</b>	<b>Unmanaged drift</b> – prompt or embedding changes silently alter results	<ul style="list-style-type: none"> <li>• <b>Info Storage</b> – Log all queries &amp; responses; keep 13-month retention</li> <li>• <b>Process Review</b> - Quarterly RAI review: bias metrics, drift tests, incident post-mortems</li> </ul>



# Future Vision

## Additional Features to Help United Take Action

1

### Proactive, Not Reactive

- **Auto-detect trends** & surface insights (e.g., food complaints spiking on Route X)
- **Real-time alerts** for emerging customer pain points

2

### Long-Term Insight

- Continue to track sentiment shifts over time as new data rolls in, by route/class
- Measure **impact of service changes & seasonal trends**

3

### Global Reach

- Include all reviews from United's 300+ international destinations
- Performant **LLM-based translation/semantic grouping** for multilingual reviews



**UNITED**

Problem

Solution

Considerations

Demo!

# Demo!





**UNITED**

**Q&A**





**UNITED**

# Appendix



# Data Source - Supplemental Detail

**Data:** A scraped dataset created from all user reviews found on Skytrax ([www.airlinequality.com](http://www.airlinequality.com))



**Goal:** Leverage scraped Skytrax dataset as a proxy for United data / review to create an initial Proof of Concept

Dataset	# of Reviews	Key Columns
Airline Reviews	41396	airline_name (object), link (object), title (object), author (object), author_country (object), date (object), content (object), aircraft (object), type_traveller (object), cabin_flow (object), route (object), overall_rating (float64), seat_comfort_rating (float64), cabin_staff_rating (float64), food_beverages_rating (float64), inflight_entertainment_rating (float64), ground_service_rating (float64), wifi_connectivity_rating (float64), value_money_rating (float64), recommended (int64)
Airport Reviews	17721	airport_name (object), link (object), title (object), author (object), author_country (object), date (object), content (object), experience_airport (object), date_visit (object), type_traveller (object), overall_rating (float64), queuing_rating (float64), terminal_cleanliness_rating (float64), terminal_seating_rating (float64), terminal_signs_rating (float64), food_beverages_rating (float64), airport_shopping_rating (float64), wifi_connectivity_rating (float64), airport_staff_rating (float64), recommended (int64)
Lounge Reviews	1258	airline_name (object), link (object), title (object), author (object), author_country (object), date (object), content (object), lounge_name (object), airport (object), lounge_type (object), date_visit (object), type_traveller (object), overall_rating (float64), comfort_rating (int64), cleanliness_rating (int64), bar_beverages_rating (float64), catering_rating (float64), washrooms_rating (float64), wifi_connectivity_rating (float64), staff_service_rating (float64), recommended (int64)
Seat Reviews	2264	airline_name (object), link (object), title (object), author (object), author_country (object), date (object), content (object), aircraft (object), seat_layout (object), date_flow (object), cabin_flow (object), type_traveller (object), overall_rating (float64), seat_legroom_rating (int64), seat_recline_rating (int64), seat_width_rating (int64), aisle_space_rating (int64), viewing_tv_rating (float64), power_supply_rating (float64), seat_storage_rating (float64), recommended (int64)

