

DialogueDNA

Emotion-Aware Speaker Analysis & Summarization



Midterm Project Presentation

Amal Amsis · Yarden Daniel · Yaniv Zamir

The Academic College of Tel Aviv-Yaffo



AGENDA

Introduction

Why DialogueDNA?

Application Development

- Use Cases
- System Flow
- AI Models & Technology Stack

UI Design

Interface, user input, and output flow

Demo

A walkthrough of a real example session

Summing Up

- Current Status
- Key Takeaways
- Open Questions

Why DialogueDNA ?

The Problem

In group conversations - such as lessons, meetings, or therapy session - it's often hard to know:

- Who said what
- How each person felt
- What triggered emotional shifts

What Exists Today

There are powerful tools available:

- **Transcription:** Whisper, AssemblyAI, Azure Speech-to-Text
- **Emotion Classification:** Hugging Face, Affectiva, Microsoft Azure Emotion API
- **Summarization:** GPT, Google T5

However, no existing tool combines all three together into single, seamless solution.

Our Goal & Solution

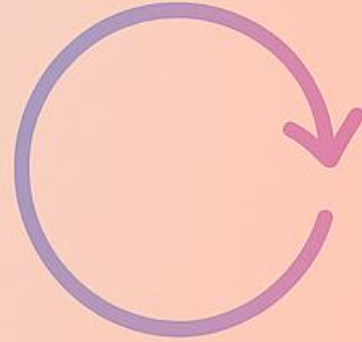
Goal:

Leverage GenAI transform raw conversations into emotion-aware, structured summaries — helping people reflect, learn, and communicate more effectively.

Our Solution Combine:

- Speaker identification
- Emotion detection (per speaker)
- Personalized summaries for each participant

APPLICATION DEVELOPMENT



Use Case Scenarios

our system adapts use case by generating emotion-awareness summaries tailored to situation and audience.



Education

Understand student participation and emotional engagement in class discussions



Therapy & Counseling

Track emotional journeys and triggers in group or individual therapy sessions



Business Meetings & Feedback

Help managers and teams reflect on emotional responses and communication dynamics



Customer Support

Evaluate how clients feel during service interactions, and how agents respond emotionally



Market Research & Focus Groups

Reveal participants' emotional reactions to products, ideas, or campaigns



Personal Development

Provide individuals with emotionally insightful summaries of important conversations

Tailored AI Summarization – Prompt Styles

Use Case	Prompt Style	Summary Focus
🧑‍💻 Customer Support	customer_service_summary	Emotional journey, agent evaluation, retention insight
🧠 Therapy & Counseling	clinical_summary / per_speaker_summary	Speaker-specific emotions, shifts, psychological patterns
💼 Team Meetings	analytical_report	Emotional distribution, trends, group dynamics
🗣️ Personal Stories	emotional_story	Emotional nuance, connection, authenticity
🔗 Combined Use	all_in_one	Narrative, emotional flow, per-speaker insights – all in one

Real Example: Using Customer Service Prompt

Session Summary

1. Interaction Summary

The customer contacted support after an extended hold time of one hour, expressing intense frustration and anger. Their primary concern was to resolve an unexplained extra fee charged to their account. The agent apologized, gathered information, and promised to investigate and follow up via the customer's preferred channel (text message). The call concluded with the agent reiterating their commitment to provide an update within a couple of hours, but the customer remained dissatisfied, expressing skepticism based on past experiences.

2. Customer Emotional Journey

The interaction began with **strong anger and frustration** due to the prolonged hold time: **"I have been on hold for an hour and I'm so angry right now."**

After voicing their issue, the customer's tone shifted to neutral but remained terse, signaling impatience and a desire for resolution.

The agent's apologies and offers to help were met with guarded responses; the customer was cooperative but clearly **remained on edge and unconvinced**.

When the agent needed contact details, the customer responded curtly, indicating mild irritation at being asked for information they believed the company already possessed.

The **emotional low point**: The customer's skepticism peaked with **"I've heard that before."** and **"Call center very angry."**—signifying deep-seated frustration and distrust.

Throughout, the customer did not display any meaningful relief or satisfaction, and their final emotional state was **unresolved anger and skepticism**.

3. Emotional Triggers & Causes

Negative Triggers:

Extended hold time: Immediate and intense anger at the outset, setting a hostile tone.

Perceived inefficiency: Frustration resurfaced when the agent requested contact information the customer felt should be easily accessible.

Unresolved history: The customer's **"I've heard that before."** comment points to unmet expectations in previous interactions, amplifying distrust.

Positive Triggers:

The agent's consistent apologies and clear communication offered minor de-escalation, but never fully alleviated the customer's frustration.

4. Agent Performance Evaluation

Strengths:

The agent maintained professionalism, responded promptly, and expressed empathy through apologies and reassuring statements.

They offered to communicate via the customer's preferred channel and set clear expectations for follow-up.

Areas for Improvement:

Empathy Depth: While apologies were present, the agent could have personalized their response further to acknowledge the customer's unique frustration, especially regarding the extensive hold time and past negative experiences.

Proactive Ownership: The agent could have reassured the customer more strongly about the seriousness of their complaint, possibly by escalating the issue or offering a token of goodwill.

Resolution Confidence: The customer's skepticism suggests the agent's assurances did not instill confidence—more assertive language or immediate partial resolution might have helped.

5. Recommendations for Improvement

Empathetic Acknowledgment: Begin with a direct acknowledgment of the customer's emotional state: "I can imagine how frustrating it must be to wait so long, and I'm truly sorry for that experience."

Ownership & Escalation: Offer to escalate cases where customers express repeated dissatisfaction or distrust.

Reduce Redundancy: Avoid asking for information the customer believes you already have; instead, confirm details on file and only request updates if necessary.

Proactive Goodwill: Where appropriate, consider offering a gesture of goodwill (e.g., fee waiver, small credit) to demonstrate the company values the customer's time and loyalty.

Empathy Training: Provide agents with ongoing training focused on emotional intelligence and handling escalated customers.

6. Conclusion

Practical Outcome: The customer's issue was not resolved during the interaction; a follow-up was promised.

Emotional Outcome: The customer left the conversation still feeling angry, skeptical, and undervalued.

Follow-up: Strongly recommended. Ensure the promised update is delivered promptly, and consider a personal call from a supervisor or manager to reinforce commitment to customer care.

7. Optional: Customer Retention Insight

Customer Sentiment: The customer currently feels alienated and distrustful toward the brand. Their repeated negative experiences and skepticism suggest a high risk of churn.

Likelihood to Return or Recommend: Low, unless the follow-up is handled exceptionally well and their issue is resolved satisfactorily.

Suggested Gesture: A direct apology from a manager, along with a compensation offer (such as refunding the extra fee or providing a service credit), could help rebuild trust and demonstrate the company's commitment to customer satisfaction.

Summary:

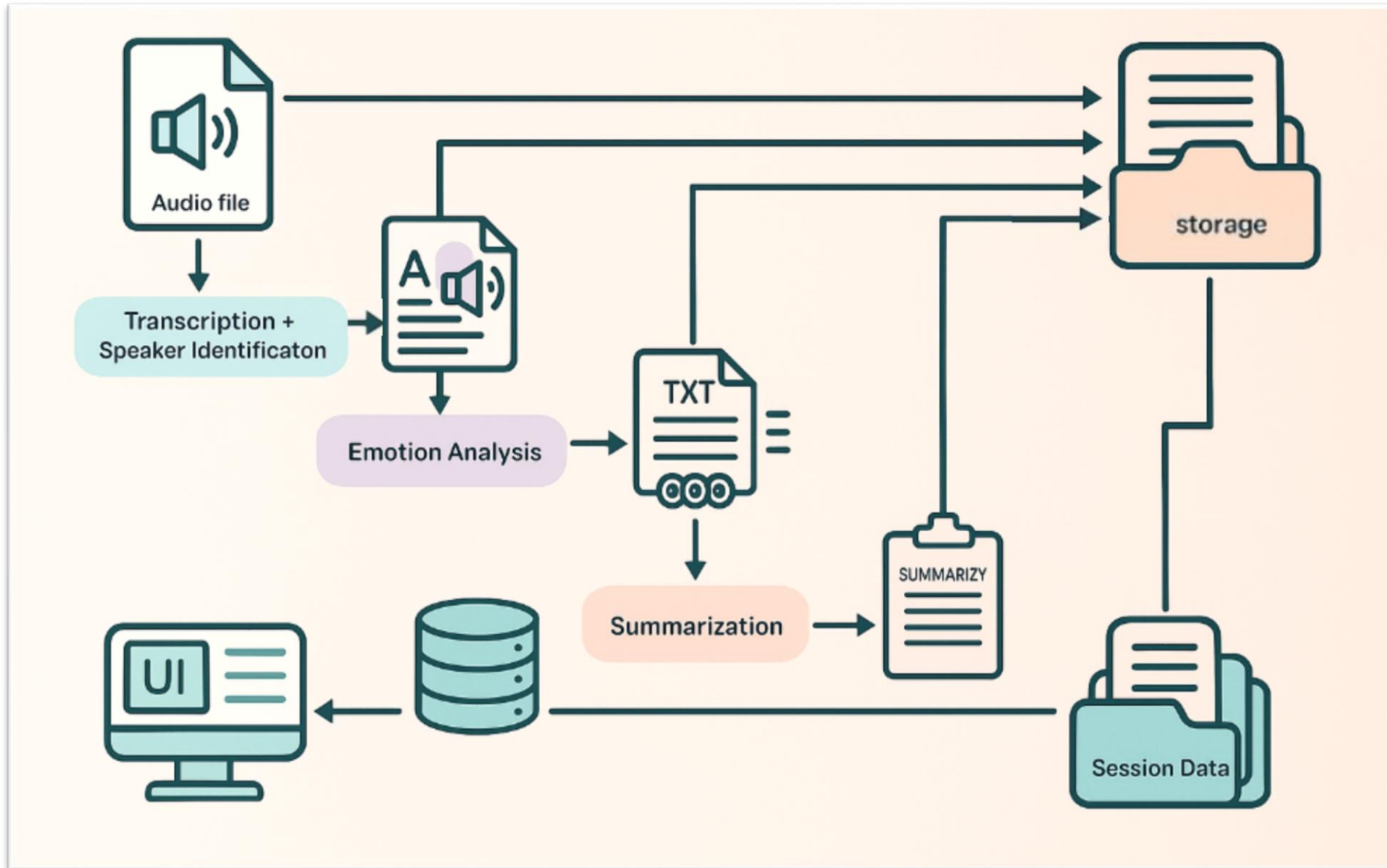
This interaction highlights the critical importance of emotional intelligence and swift, personalized responses—especially for customers with a history of negative experiences. By addressing not just the practical issue but also the underlying emotional distress, future interactions can improve both customer satisfaction and loyalty.

Customer Service Prompt

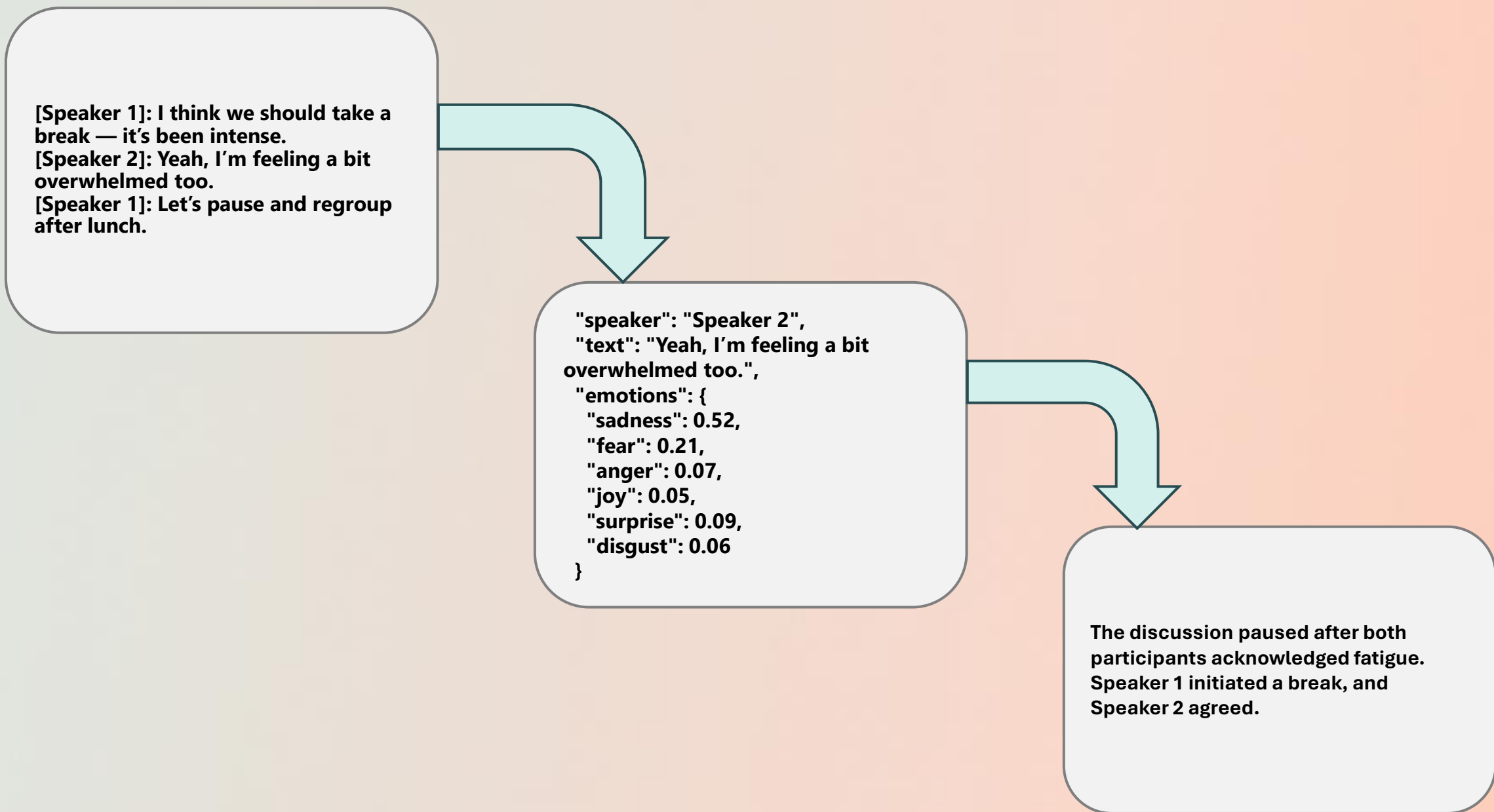
```
"customer_service_summary": (  
  "You are a professional customer experience analyst with expertise in emotional intelligence and conversation behavior.\n"  
  "You've received a transcript of a service interaction between a customer and a support agent. The transcript includes speaker labels and emotional annotations.\n"  
  "Your task is to write a clear, structured, and emotionally insightful summary of this interaction.\n"  
  "Focus on the customer's emotional journey, identify key emotional triggers, evaluate the agent's performance, and offer recommendations for improvement.\n"  
  "Avoid quoting raw emotion scores - translate them into meaningful human interpretations.\n"  
  "Write in a professional yet compassionate tone.\n\n"  
  "Structure your output with the following sections:\n\n"  
  "📌 1. Interaction Summary\n"  
  "- What was the customer's issue or request?\n"  
  "- What actions were taken and what was the final outcome?\n\n"  
  "🌀 2. Customer Emotional Journey\n"  
  "- How did the customer feel during the interaction?\n"  
  "- Identify emotional turning points.\n"  
  "- Use bold for emotionally significant lines or reactions.\n"  
  "- Reflect on whether the customer felt heard and understood.\n\n"  
  "🚩 3. Emotional Triggers & Causes\n"  
  "- What caused any negative or positive emotional shifts?\n"  
  "- Be specific about moments that escalated or de-escalated tension.\n\n"
```

```
"👤 4. Agent Performance Evaluation\n"  
  "- How well did the agent respond emotionally and professionally?\n"  
  "- What worked well, and what could have been improved?\n"  
  "- Focus on empathy, clarity, tone, and resolution.\n\n"  
  "🔧 5. Recommendations for Improvement\n"  
  "- Offer concrete suggestions to improve future service experiences.\n"  
  "- These can include phrasing changes, empathy training, or process adjustments.\n\n"  
  "🏆 6. Conclusion\n"  
  "- Was the issue resolved practically and emotionally?\n"  
  "- What emotional state did the customer leave in?\n"  
  "- Is follow-up recommended?\n\n"  
  "📌 7. Optional: Customer Retention Insight\n"  
  "- Based on the conversation, what is the customer likely to feel toward the brand?\n"  
  "- Would they return, churn, or recommend the service?\n"  
  "- Suggest a possible gesture (e.g., apology, compensation) if appropriate.\n\n"  
  "Be detailed, empathetic, and focused on delivering insights that can improve both the agent's performance and the overall customer experience."
```


Application Flow



[Speaker 1]: I think we should take a break — it's been intense.
[Speaker 2]: Yeah, I'm feeling a bit overwhelmed too.
[Speaker 1]: Let's pause and regroup after lunch.



```
graph LR; A["[Speaker 1]: I think we should take a break — it's been intense.  
[Speaker 2]: Yeah, I'm feeling a bit overwhelmed too.  
[Speaker 1]: Let's pause and regroup after lunch."] --> B["speaker: Speaker 2,  
text: Yeah, I'm feeling a bit overwhelmed too.,  
emotions: {  
  sadness: 0.52,  
  fear: 0.21,  
  anger: 0.07,  
  joy: 0.05,  
  surprise: 0.09,  
  disgust: 0.06  
}"]; B --> C["The discussion paused after both participants acknowledged fatigue. Speaker 1 initiated a break, and Speaker 2 agreed."];
```

"speaker": "Speaker 2",
"text": "Yeah, I'm feeling a bit overwhelmed too.",
"emotions": {
 "sadness": 0.52,
 "fear": 0.21,
 "anger": 0.07,
 "joy": 0.05,
 "surprise": 0.09,
 "disgust": 0.06
}

The discussion paused after both participants acknowledged fatigue. Speaker 1 initiated a break, and Speaker 2 agreed.

Behind the Scenes: GenAI Logic

The key models powering transcription, emotion detection, and summarization

Component	Model Used	Role in the System
Transcription	Microsoft Azure Speech Service (Conversation Transcription), incling Speaker identification	• Speaker identification & timestamped ASR • Multi-speaker support conversations. • Delivers structured output to the emotion detection layer.
Emotion Detection	Hugging Face – j-hartmann/emotion-english-distilroberta-base	• Analyzes text utterance-level emotion classification. • Identifies six emotions: joy, sadness, anger, fear, surprise, disgust. • Fine-tuned on dialogue-style data
Summarization	GPT-4.1 (via Azure OpenAI)	• Generates speaker-aware, context-rich summaries. • Uses conversation-type-specific prompts (e.g., therapy vs. Meeting) • Supports emotional/contextual nuance

Future Plan: Integrate voice-tone-based emotion analysis using acoustic models

DialogueDNA — Technology Stack



AI & NLP Models

Azure OpenAI – GPT-4.1

→ Summarization

Hugging Face –

DistilRoBERTa →

Emotion Detection

Azure Speech-to-Text

(with Speaker
Identification) →

Transcription



Backend

Python

FastAPI → API endpoints
and logic coordination



Cloud & Storage

Azure Blob Storage →

Audio, transcript,
emotion & summary files

Supabase / PostgreSQL

→ Session metadata &
state

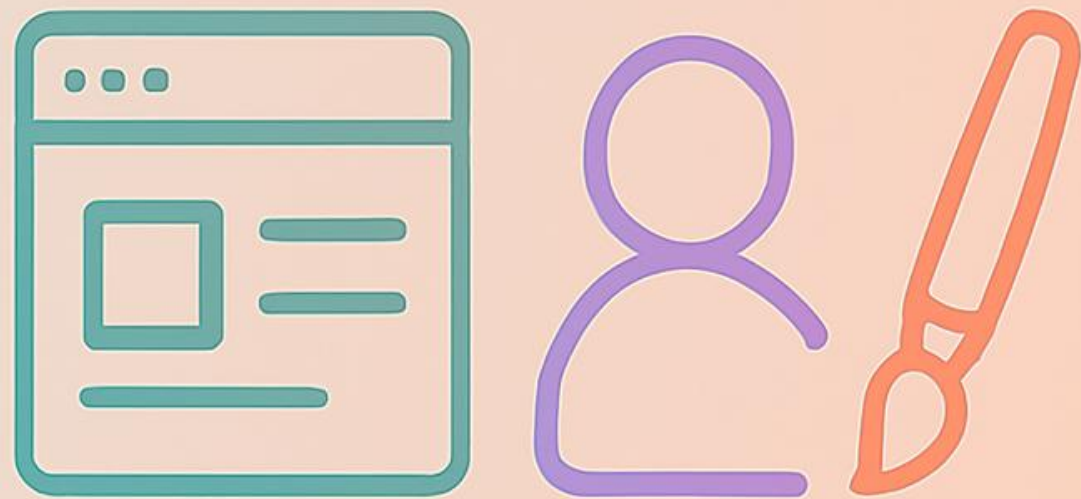


Frontend

React → UI

React Query → Data
syncing & state

UI DESIGN



Settings

- **Login Page:**
 - Input fields for **Email** and **Password**
 - *Soon: "Forgot password?" link*
- **Sign Up Page:**
 - Input fields for **Full Name**, **Email**, **Password**, and **Confirm Password**
- **Authentication:**
 - Handled via **Supabase**
 - Session management and token-based security
 - *Soon: Supports **Google** and **Facebook** login options*



EmotionAI

Email

Password

Login

Don't have an account? [Sign Up](#)

User Input

- **Audio Upload Interface**
 - Drag-and-drop area or file upload button
 - Accepts .wav
 - *Soon: Accept all types of audio*
- **Soon: Session Settings**
 - *Language* selection dropdown
 - Audio **Summarization** Type
 - *Speaker count* input (optional)
- **File Validation**
 - Displays **file format**, **size limits**, and **errors** if invalid
- **Upload Status**
 - Clear loading indicator with appropriate text.
 - *Soon: Real-time progress bar*

Welcome to EmotionAI Tool

You haven't created any sessions yet.

Start Your First Session

Create New Session ×

Session Title

Upload Audio File

Choose File No file chosen



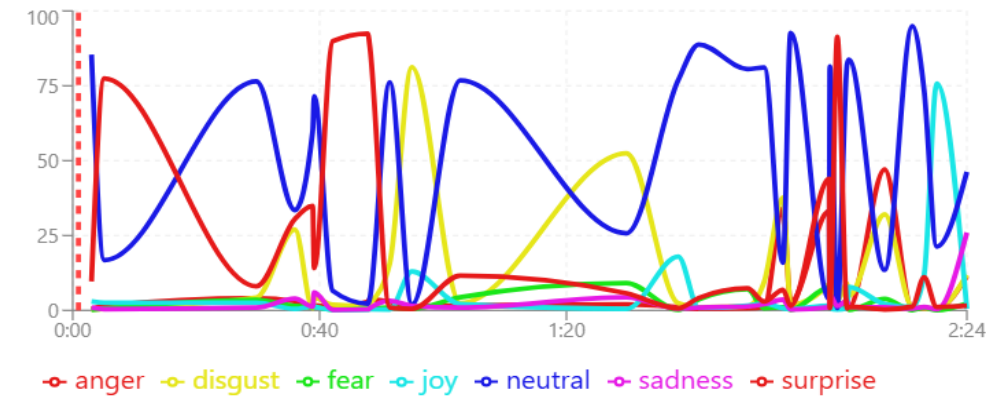
Cancel

Create Session

Application Output

- **Transcription View:**
 - Full transcript with **speaker labels**
- **Emotion Analysis:**
 - Emotion tags per **speaker**
 - X-Axis: Audio **time** from start to end
 - Y-Axis: **Intensity level** of each emotion
- **Summary Generator:**
 - AI-generated summaries tailored to context (e.g., emotional, analytical)
- **Soon: Export Options:**
 - *Download emotional analysis, transcript and summary as **PDF**.*

Speaker 1 - Emotional Analysis 🎤



Transcript

Speaker 1:

You made two Wonder Woman movies since the last time I'm.

Speaker 2:

So exhausted.

Speaker 1:

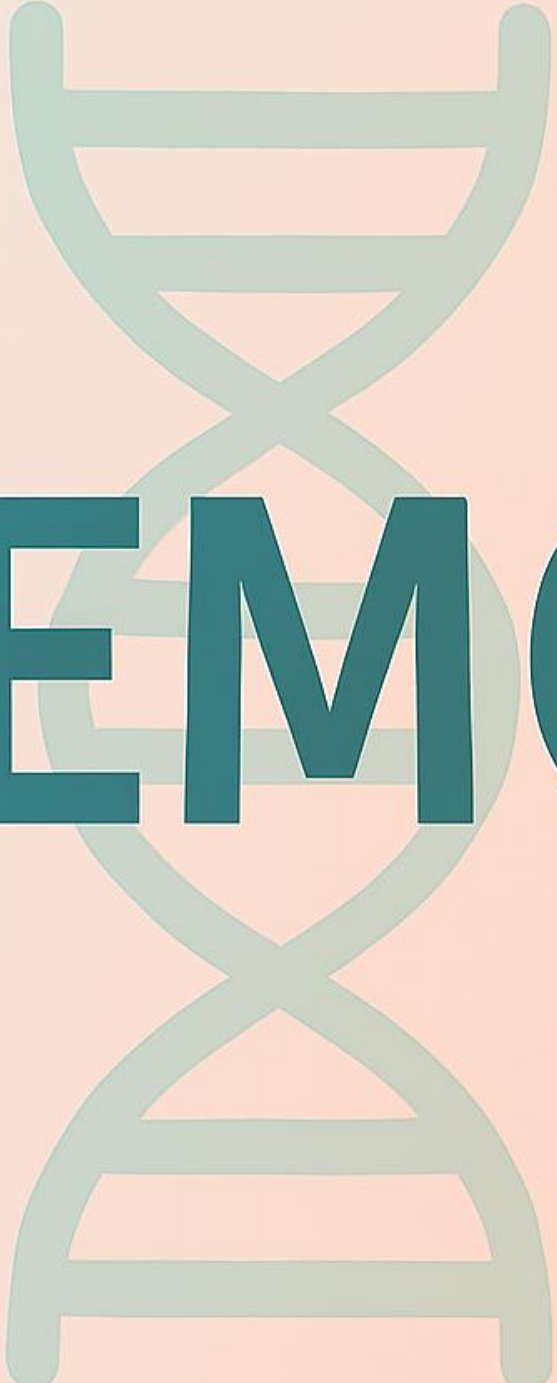
Wonder Woman, are you?

Session Summary

📖 Opening Lines: Catching Up on Superheroes

The conversation begins with a gentle, almost bemused acknowledgment of time's passage:

"You made two Wonder Woman movies since the last time I'm." There's a sense of admiration here, but also a subtle awe at the sheer volume of achievement. It sets the tone for a dialogue that will oscillate between playful banter and genuine reflection on exhaustion, pride, and family.



DEMO



EmotionAI

Email

Password

Login

Don't have an account? [Sign Up](#)

RESULT



GENAI REQUIREMENT : BENCHMARK RESULTS

Component	Metric Evaluated	Test Method	Acceptance Criteria	Result
Transcription	Speaker Identification Accuracy	Manual comparison of speaker	≥ 90% correct speaker assignment	✅ Azure STT passed (~93%)
	Transcription Quality	Manual reading of transcripts from the same clips	Fluent sentence structure, few omissions	✅ Azure STT passed
	Overlap	Tested on short clips where speakers overlapped	correct speaker assignment And Fluent sentence structure, few omissions	✅ Azure STT handled overlap best
Emotion Detection	Emotion Label Accuracy	Compared model predictions to manual labels on known-emotion clips	≥ 80% agreement with human-labeled emotions	✅ Hugging Face passed (~85%)
Summarization	Prompt Adaptability	Iterative testing of multiple prompt styles for different user needs	Summary tone and content adjusted to prompt type	✅ GPT-4 passed
	Clarity & Personal Relevance	Reviewed summaries for different speakers in real sessions	Covers key content and reflects emotional tone	✅ GPT-4 passed

**SUMMING
UP**

Status

What has been done?

- **Backend pipeline established**
Upload → Transcribe → Emotion analysis
→ Summary
- **User interface developed**
Fetch and display session results.
- **Session management**
Implement via Azure Blob and Supabase.
- **Emotion-aware summaries**
Structured and speaker-specific.

SOON What will be done?

- **Tone-based emotion detection**
Voice emotion detection and combine with text analysis for richer insights.
- **Refine UI** : For improve clarity, flow and experience
- **Support live mode**: For real-time conversations, not just uploads.

*“I’ve learned that people will forget what you said,
people will forget what you did, but people will
never forget how you made them feel.”*

— Maya Angelou

Thank you

Amal Amsis · Yarden Daniel · Yaniv-Zamir

The Academic College of Tel Aviv-Yaffo