

Popular Generative AI Tools (Application-wise)

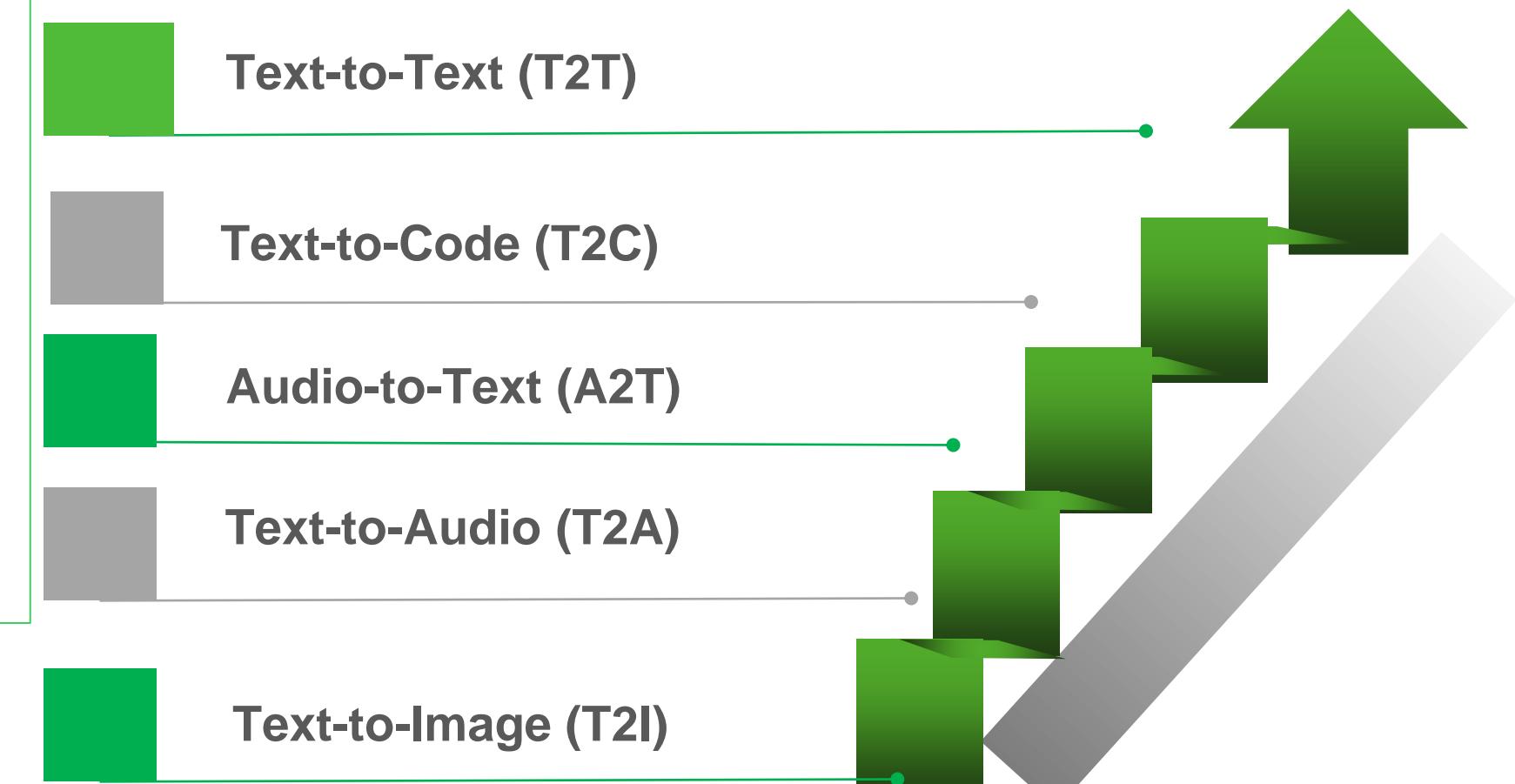
Speaker : Shubham Bhardwaj

AI/ML Technical Lead

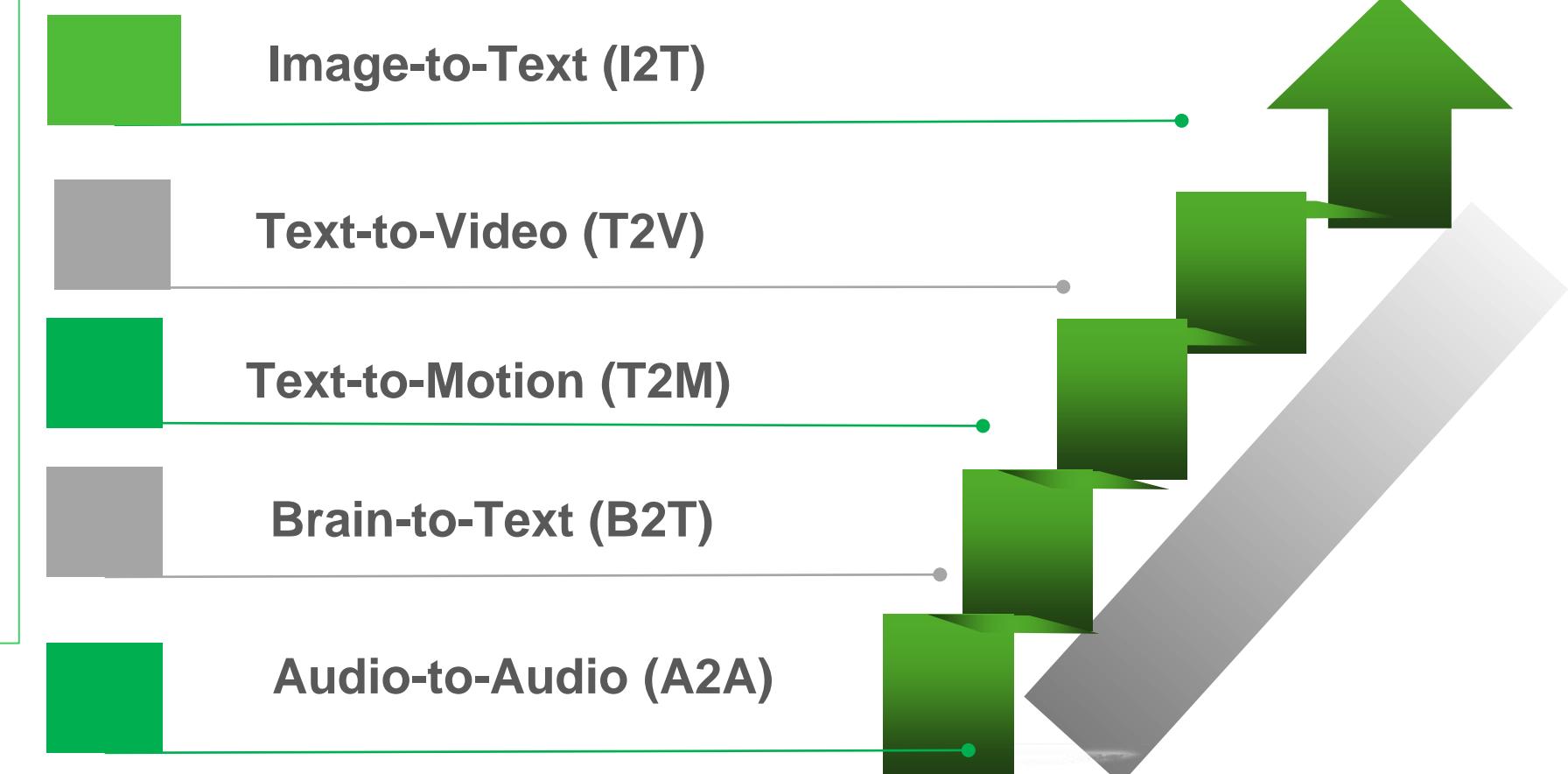
WalkingTree Technologies

<https://walkingtree.tech/>

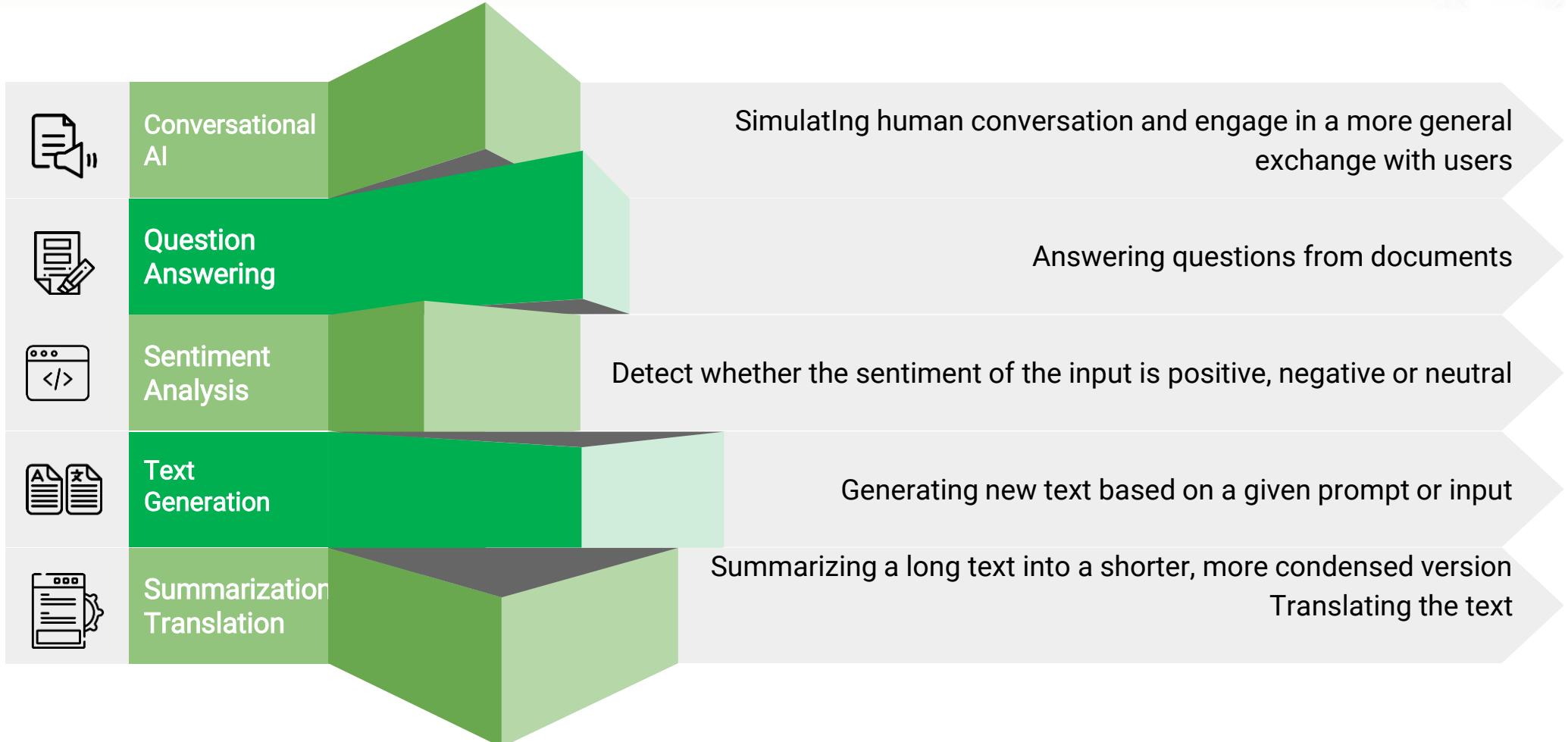
“Generative AI is a type of artificial intelligence that focuses on creating new and unique outputs from a set of input data. It is based on the idea of machine learning, where a model is trained on a large amount of data and then generates new content based on what it has learned. “



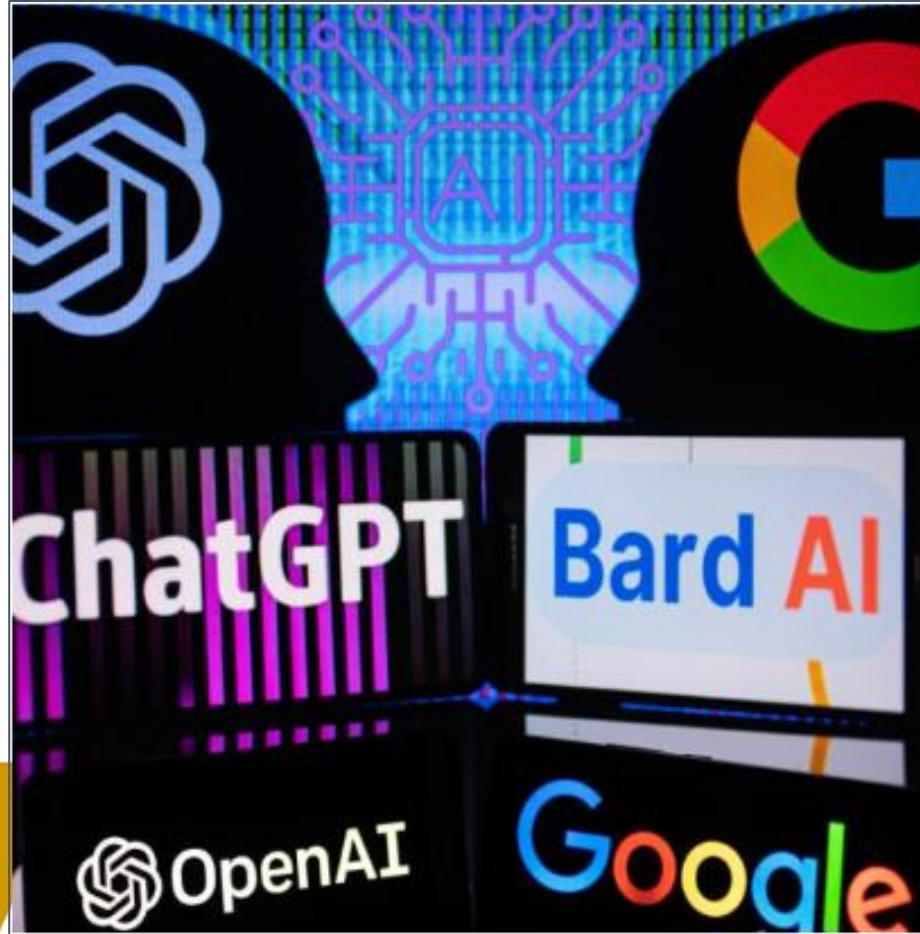
"This can help businesses and professionals to create more diverse and interesting content that appeals to a wider range of people, can be personalised and is of better quality."



Text-to-Text (T2T)



28 Feb – 04 March 2023 | New Delhi

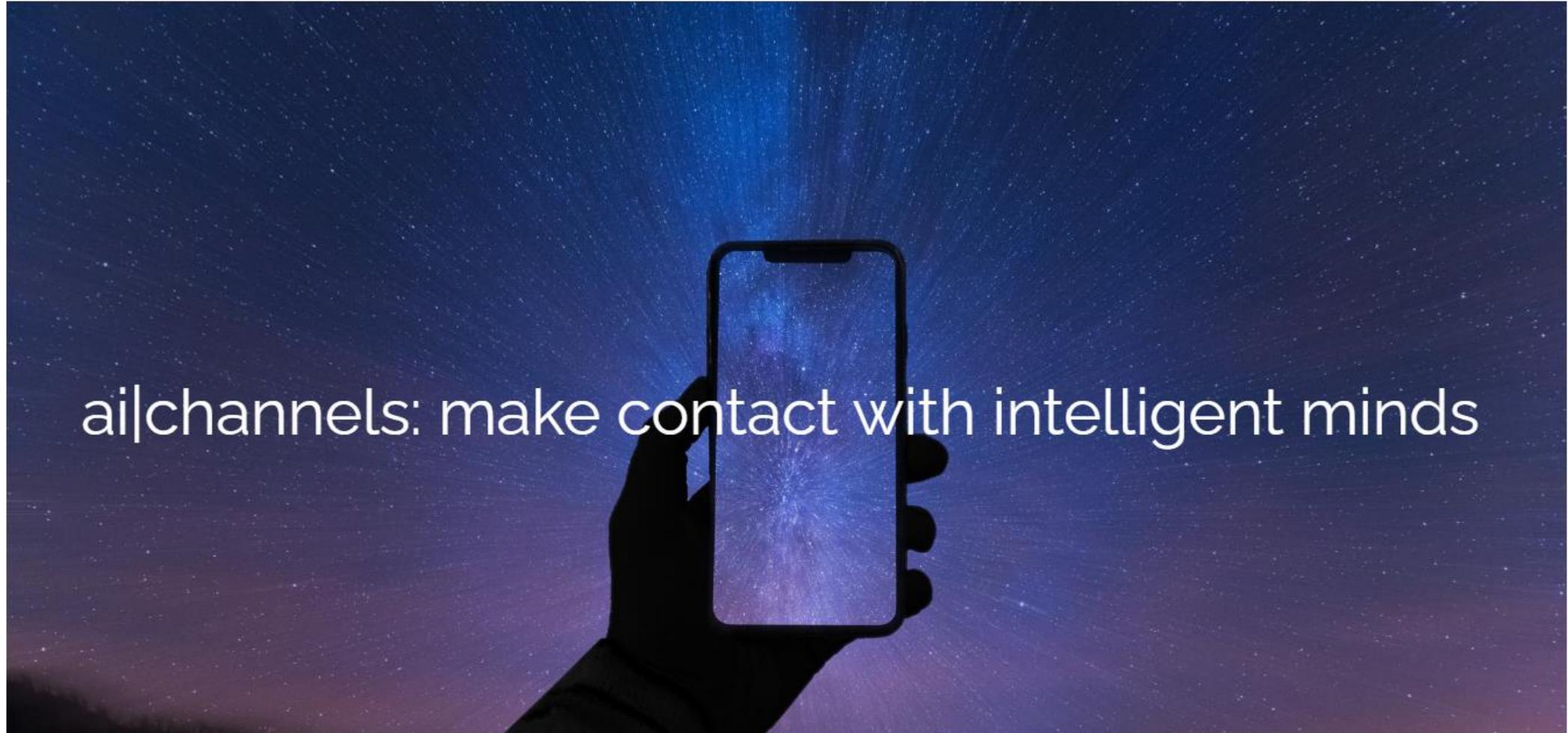


	Bard	ChatGPT
Language model	LaMDA	GPT-3
Source of Information	Internet	Data feed
Information cutoff	None	2021
Access	Limited	Unlimited
Limitations	Biases of Internet	Biases of data

Conversational AI - AI Channels

ORGANIZER

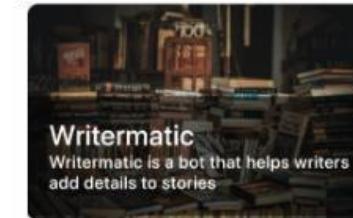
ISGF
India Smart Grid Forum



Conversational AI - AI Channels

ORGANIZER

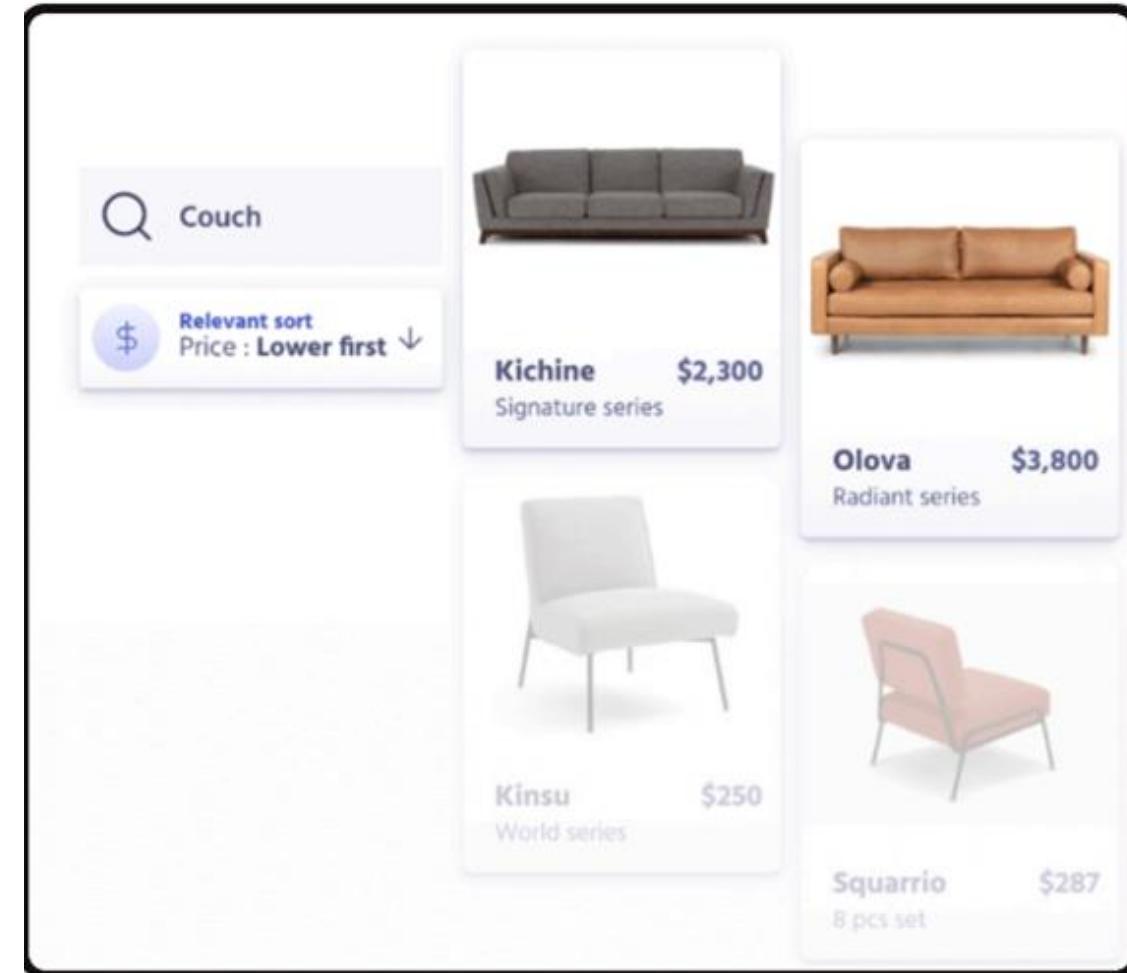
ISGF
India Smart Grid Forum



Semantic Search - Algolia

ORGANIZER

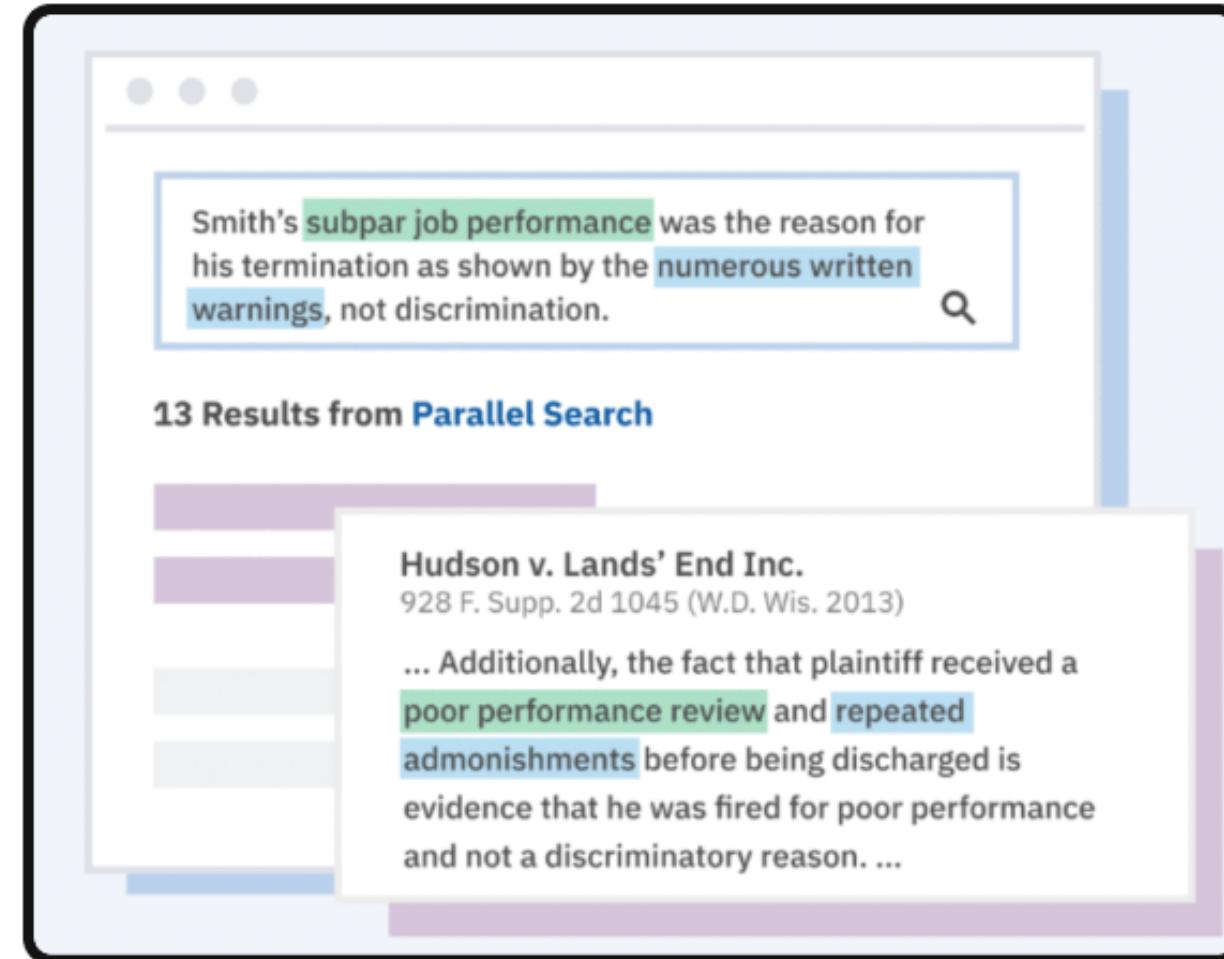
ISGF
India Smart Grid Forum



Semantic Search - Casetext

ORGANIZER

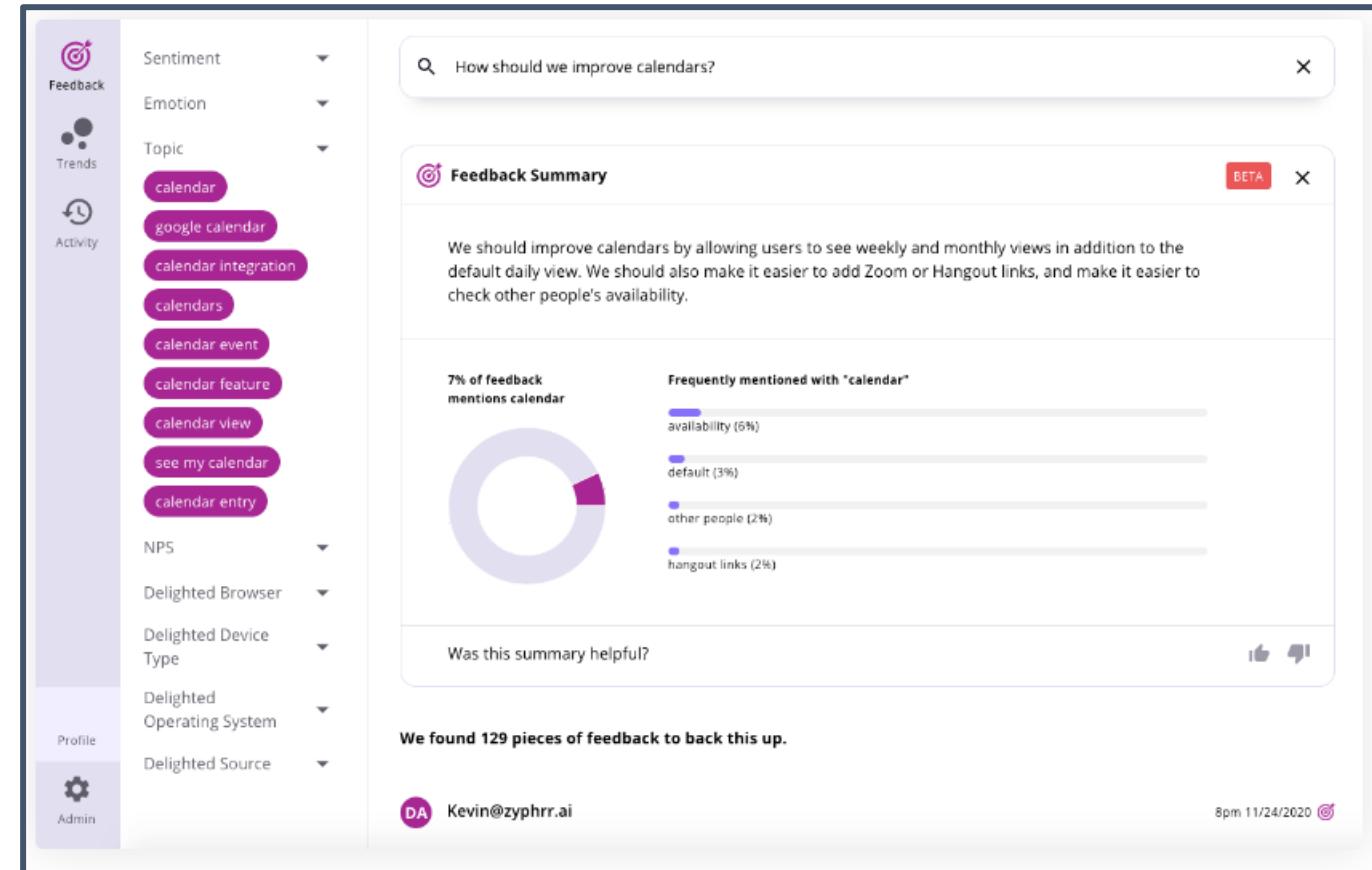
ISGF
India Smart Grid Forum



Customer Service- Viable

ORGANIZER

ISGF
India Smart Grid Forum

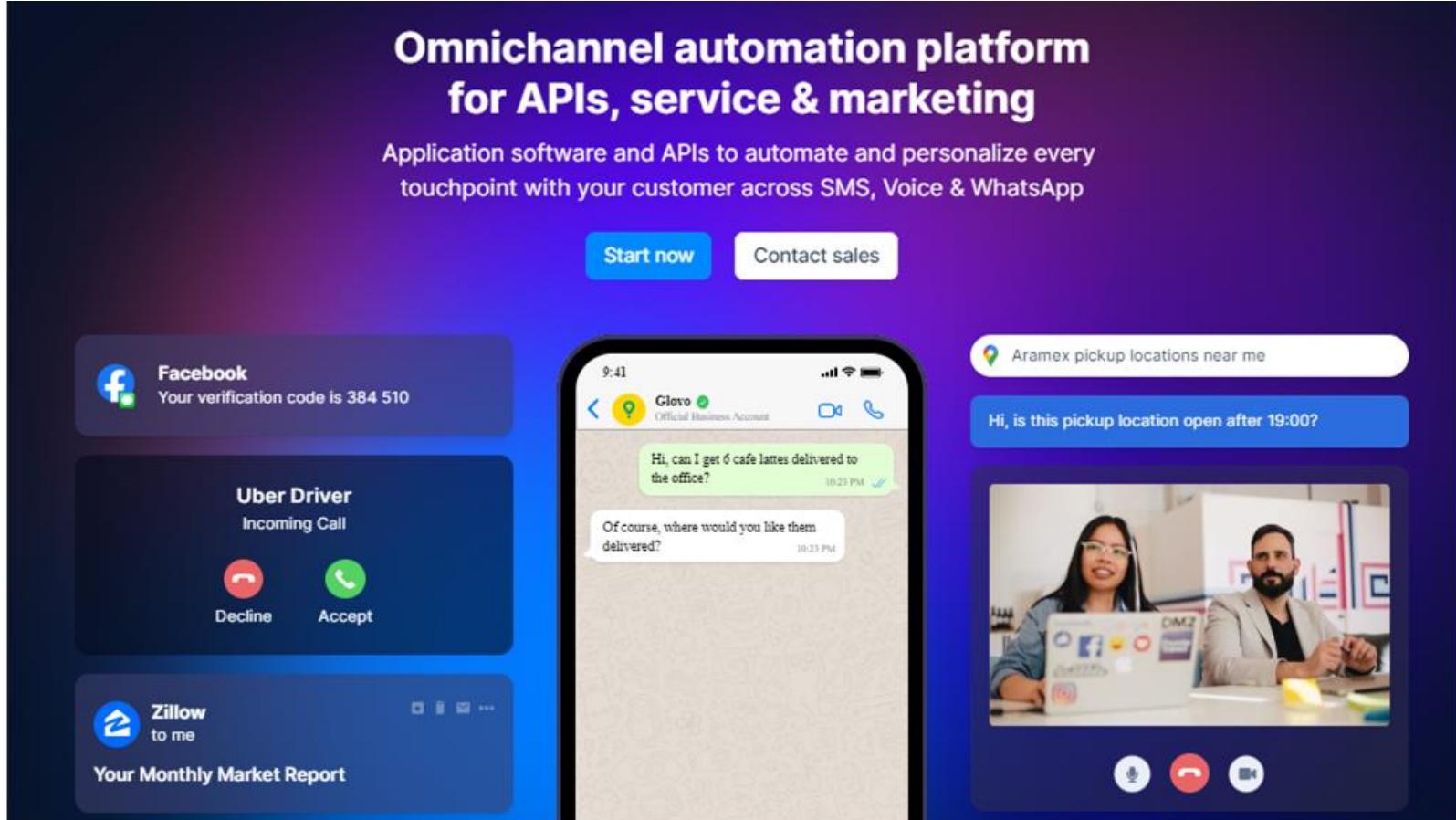


The screenshot shows a user interface for analyzing customer feedback. On the left, a sidebar lists various metrics: Sentiment, Emotion, Topic (with items like 'calendar', 'google calendar', 'calendar integration', 'calendars', 'calendar event', 'calendar feature', 'calendar view', 'see my calendar', and 'calendar entry'), NPS, Delighted Browser, Delighted Device Type, Delighted Operating System, Delighted Source, and Admin. The main area features a search bar at the top with the query 'How should we improve calendars?'. Below it is a 'Feedback Summary' section with a 'BETA' label. The summary text reads: 'We should improve calendars by allowing users to see weekly and monthly views in addition to the default daily view. We should also make it easier to add Zoom or Hangout links, and make it easier to check other people's availability.' A circular chart indicates '7% of feedback mentions calendar'. To the right, a bar chart titled 'Frequently mentioned with "calendar"' shows the following percentages: availability (6%), default (3%), other people (2%), and hangout links (2%). At the bottom, a message from 'Kevin@zyphrr.ai' is displayed with a timestamp of '8pm 11/24/2020'.

**Omnichannel automation platform
for APIs, service & marketing**

Application software and APIs to automate and personalize every touchpoint with your customer across SMS, Voice & WhatsApp

[Start now](#) [Contact sales](#)



Facebook
Your verification code is 384 510

Uber Driver
Incoming Call
Decline Accept

Zillow
to me
Your Monthly Market Report

9:41 Glovo Official Business Account

Hi, can I get 6 cafe lattes delivered to the office? 10:23 PM

Of course, where would you like them delivered? 10:23 PM

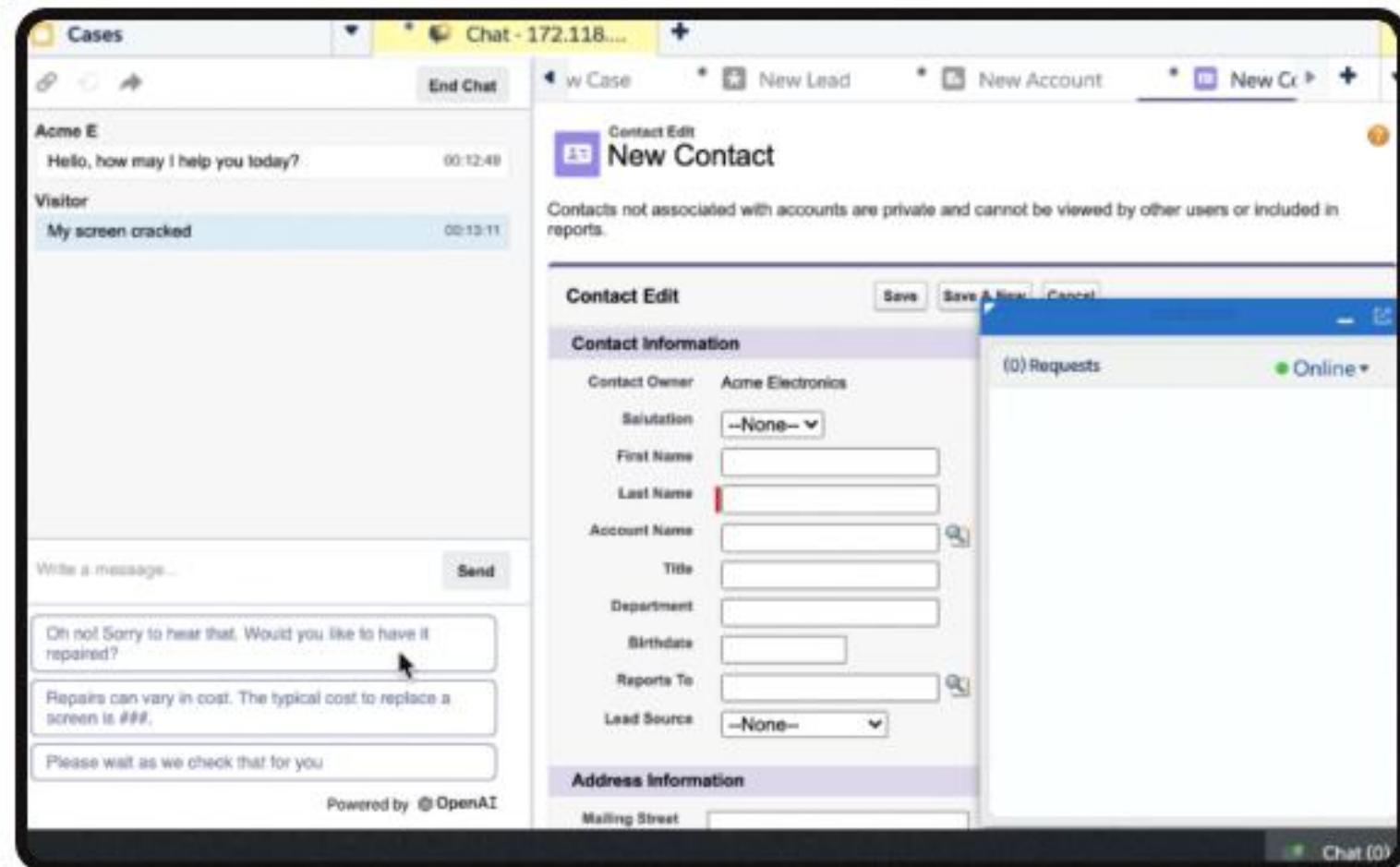
Aramex pickup locations near me

Hi, is this pickup location open after 19:00?

Customer Service- Sapling

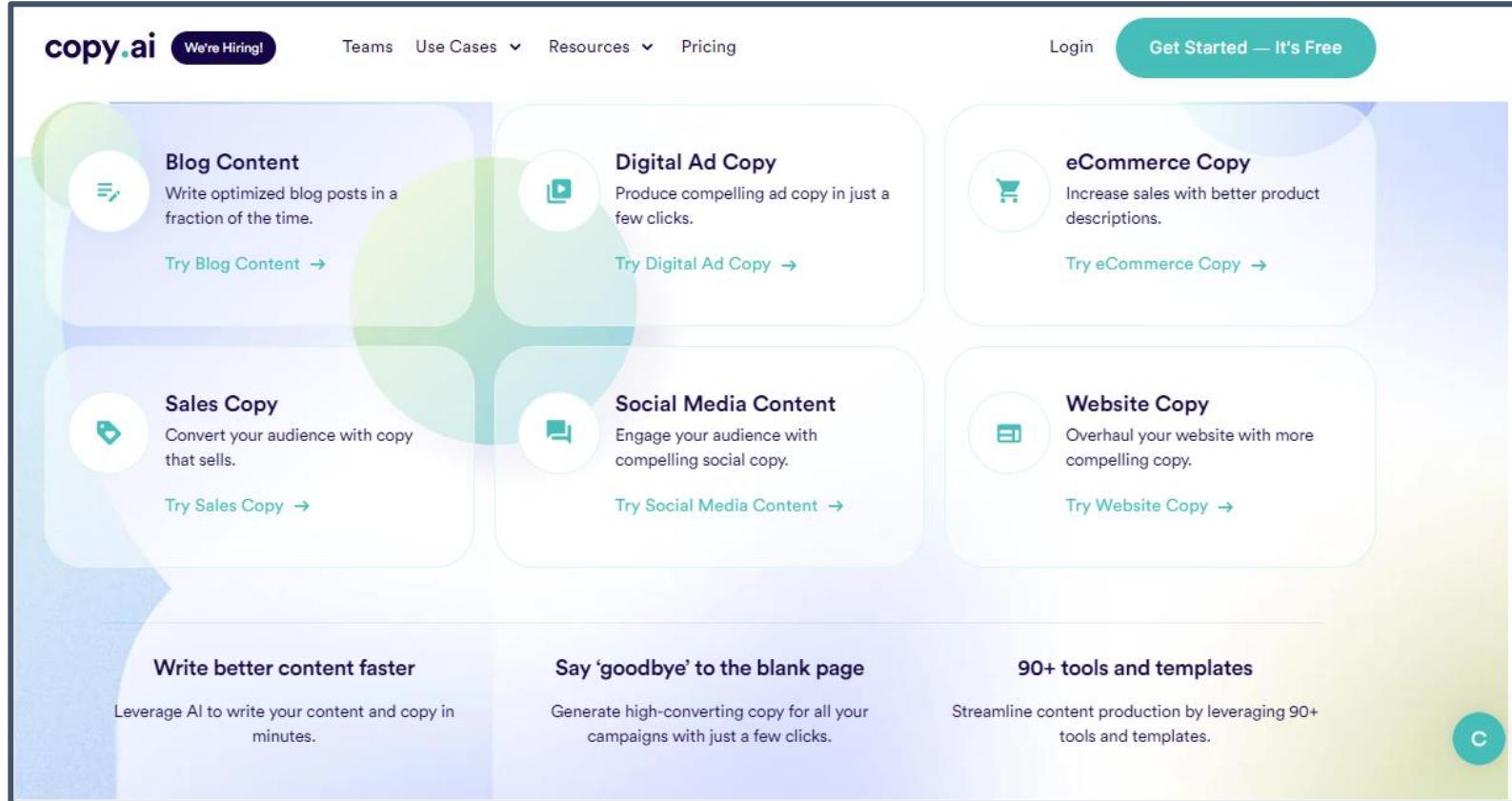
ORGANIZER

ISGF
India Smart Grid Forum



Content Creation- copy.ai

ORGANIZER



The screenshot shows the homepage of copy.ai. At the top, there's a navigation bar with the copy.ai logo, a "We're Hiring!" button, and links for Teams, Use Cases, Resources, Pricing, Login, and a prominent "Get Started — It's Free" button. Below the navigation, six cards represent different content creation tools:

- Blog Content**: Write optimized blog posts in a fraction of the time. Try [Blog Content](#).
- Digital Ad Copy**: Produce compelling ad copy in just a few clicks. Try [Digital Ad Copy](#).
- eCommerce Copy**: Increase sales with better product descriptions. Try [eCommerce Copy](#).
- Sales Copy**: Convert your audience with copy that sells. Try [Sales Copy](#).
- Social Media Content**: Engage your audience with compelling social copy. Try [Social Media Content](#).
- Website Copy**: Overhaul your website with more compelling copy. Try [Website Copy](#).

At the bottom of the page, three promotional sections highlight the platform's features:

- Write better content faster**: Leverage AI to write your content and copy in minutes.
- Say 'goodbye' to the blank page**: Generate high-converting copy for all your campaigns with just a few clicks.
- 90+ tools and templates**: Streamline content production by leveraging 90+ tools and templates.

T2T Use Cases for Energy Sector

ORGANIZER

ISGF
India Smart Grid Forum

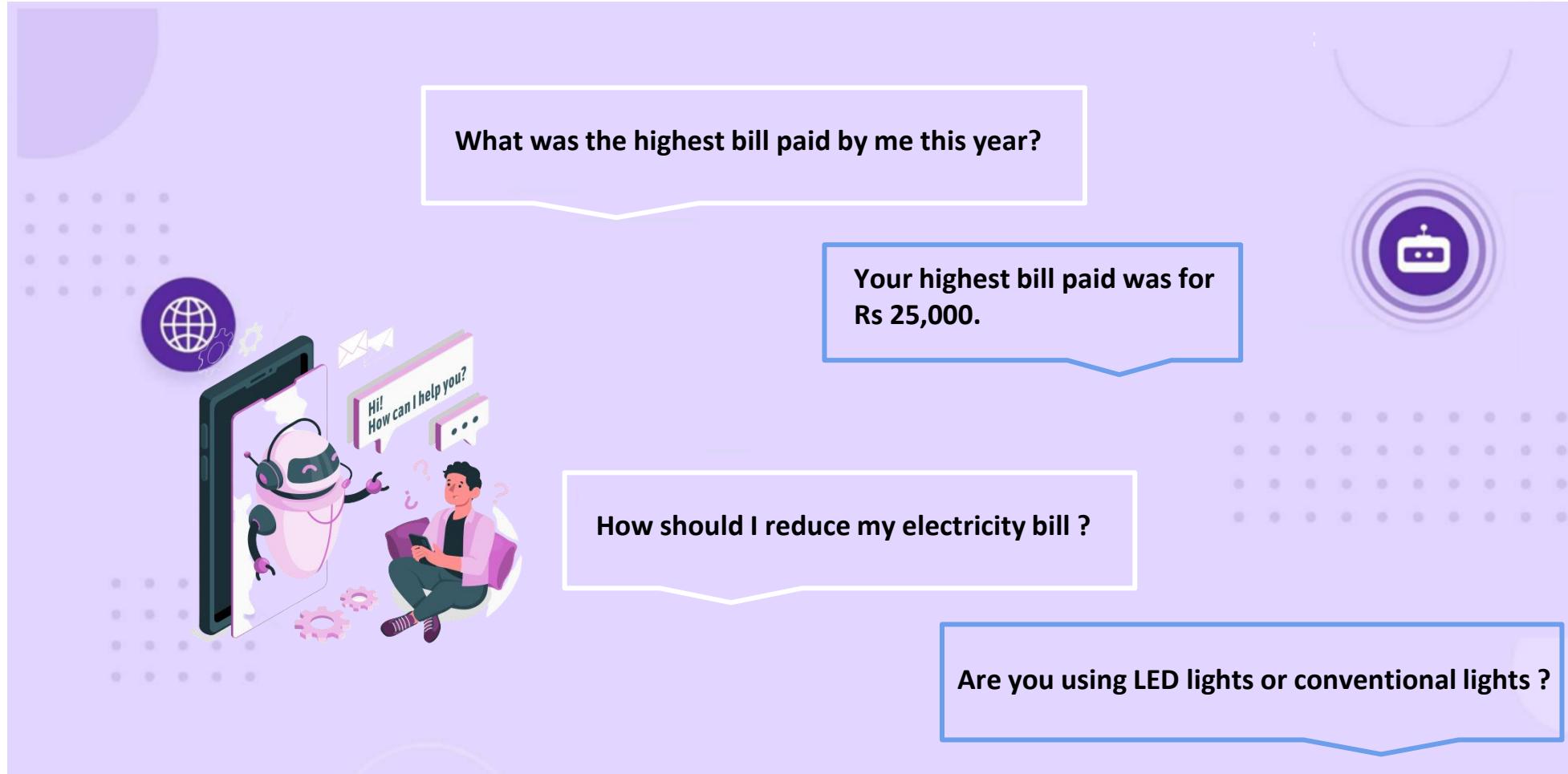
LLM powered Customer Support Chatbot

Feature	Details
Billing and Payment Inquiries	Answer questions related to billing and payment, such as providing information on account balances, due dates, and payment options
Service Outages and Maintenance	Provide information about power outages and scheduled maintenance work, including estimated restoration times and updates on the status of repairs
Account Management	Help customers manage their accounts, such as updating personal information, setting up automatic payments, and signing up for paperless billing
Technical Support	Assist customers with technical issues related to their power utility service, such as troubleshooting problems with meters, power lines, or other equipment
Energy Conservation Tips	Provide tips and advice on how to conserve energy and reduce energy costs, such as adjusting thermostat settings and using energy-efficient appliances

T2T Use Cases for Energy Sector

ORGANIZER

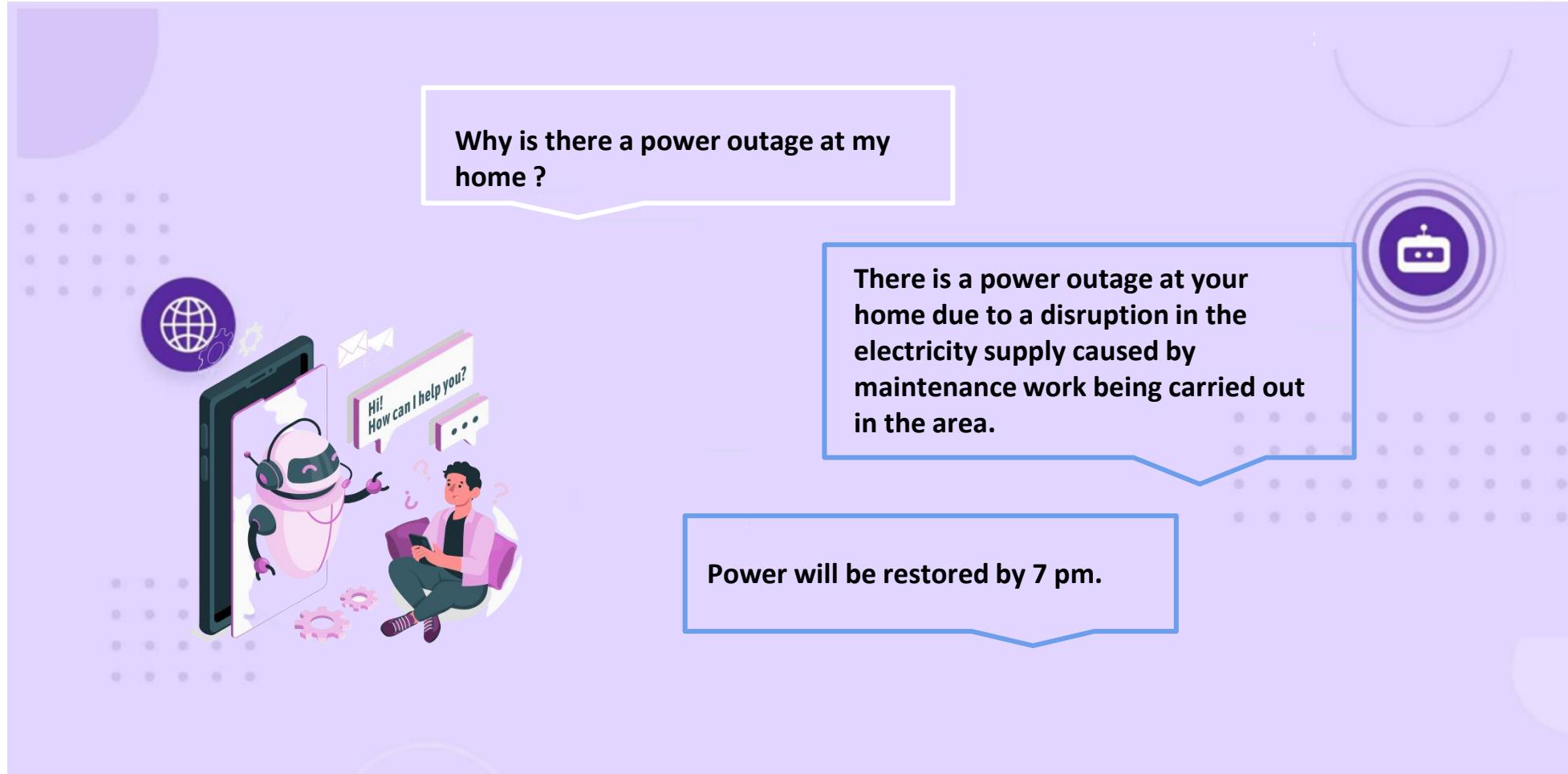
ISGF
India Smart Grid Forum



T2T Use Cases for Energy Sector

ORGANIZER

ISGF
India Smart Grid Forum



T2T Use Cases for Energy Sector

ORGANIZER

ISGF
India Smart Grid Forum

Use Case

Predictive Maintenance

Description

Predict when maintenance issues will occur

Document to Analyse

Maintenance Logs, Equipment Manuals

Safety and Risk Management

Identify potential safety issues and advise steps to prevent accidents and incidents

Safety Reports, Risk Assessments

Regulatory Compliance

Ensure compliance with legal and ethical guidelines

Regulatory documents, Industry standards

Carbon Footprint Management

Identify opportunities for reducing carbon emissions

Energy usage and Emissions documents

Text-to-Code (T2C) Models

ORGANIZER

ISGF
India Smart Grid Forum

T2C is an NLP task that involves generating computer programming codes from a given text description.

This task can be used to generate code for various programming languages, such as Python, Java, and C++. T2C is also known as "code generation" or "code synthesis".

Core Capabilities

- Fetching results from database using Natural Language
- Generate code for different kinds of programming languages
- Fixing Code vulnerabilities
- Improving code quality
- Code documentation generation

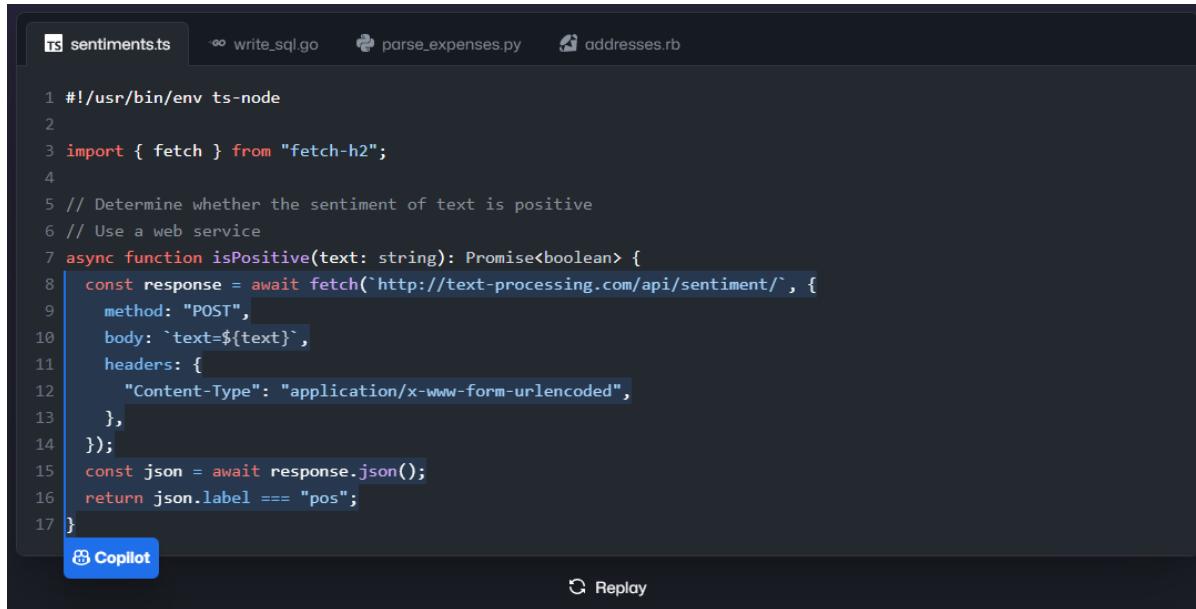
Cogram

Text-to-Code (T2C) Tools

ORGANIZER



GitHub Copilot



A screenshot of a code editor interface for GitHub Copilot. The editor shows several files: sentiments.ts, write_sql.go, parse_expenses.py, and addresses.rb. The file sentiments.ts contains the following code:

```
1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment of text is positive
6 // Use a web service
7 async function isPositive(text: string): Promise<boolean> {
8   const response = await fetch(`http://text-processing.com/api/sentiment/`, {
9     method: "POST",
10    body: `text=${text}`,
11    headers: {
12      "Content-Type": "application/x-www-form-urlencoded",
13    },
14  });
15  const json = await response.json();
16  return json.label === "pos";
17}
```

At the bottom left, there is a blue button labeled "Copilot". At the bottom right, there is a "Replay" button.

 **AI pair programmer that offers autocomplete-style suggestions as you code.**

 **You can receive suggestions from GitHub Copilot either by starting to write the code you want to use, or by writing a natural language comment describing what you want the code to do.**

Text-to-Code (T2C) Tools

ORGANIZER

ISGF
India Smart Grid Forum



```
1 # get random forest model
2 import numpy as np
3 from sklearn.model_selection import train_test_split
4 from sklearn.ensemble import RandomForestRegressor
5 from sklearn.metrics import mean_squared_error, r2_score
6
7 # load data from train.csv
8 train_df = pd.read_csv('data/train.csv')
9 train_df['target'] = train_df['target'].astype(int)
10
11 # split the data into training and testing sets
12 X_train, X_test, y_train, y_test = train_test_split(train_df, test_size=0)
13
14 # fit random forest model
15 rf = RandomForestRegressor()
16 rf.fit(X_train, y_train)
17
18 # predict labels of test set
19 y_pred = rf.predict(X_test)
20
21 # calculate mean squared error
22 mse = mean_squared_error(y_test, y_pred)
print('MSE:', mse)
```

Tabnine uses generative AI technology to predict and suggests your next lines of code based on context & syntax.

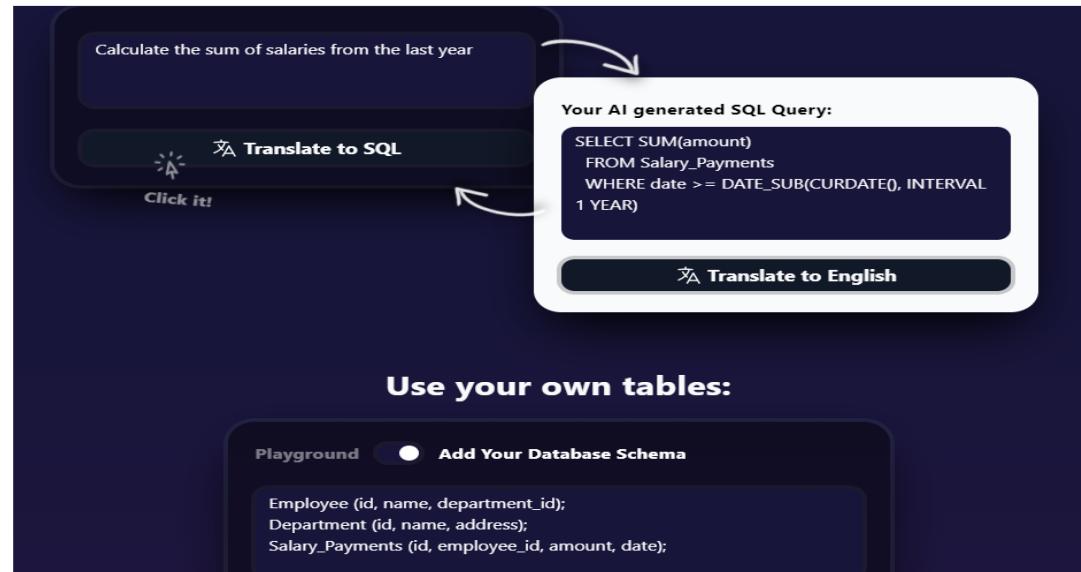
-  Whole line code completions
-  Full-function code completions
-  Natural language to code

Text-to-Code (T2C) Tools

ORGANIZER

ISGF
India Smart Grid Forum

Text to SQL



Translate English prompts to SQL queries

Translate SQL queries to English text.

Text-to-Code (T2C) Tools

ORGANIZER

ISGF
India Smart Grid Forum



debuild.co

Describe your app.

Just describe your app!

Add \$3 Withdraw \$5

Give away all my money

My balance is 7

```
// a button that says "Add $3" and
// a button that says "Withdraw $5"
// and a button that says "Give away
// all my money". then show me my
// balance
class App extends React.Component
{
```

- Lets you build web applications with simple English sentences
- Shows the code used to create the desired site environment, hence users can replicate the code in any coding platform they use.

1. Power Plants table:

This table stores information about the power plants in the power grid system, such as plant ID, plant name, location, capacity, fuel type, and status.

Plant ID	Plant Name	Location	Capacity	Fuel Type	Status
1	ABC Plant	City 1	100 MW	Natural Gas	Active
2	XYZ Plant	City 2	200 MW	Coal	Inactive
3	PQR Plant	City 3	150 MW	Nuclear	Active

2. Transmission Lines table:

This table stores information about the transmission lines that connect the power plants to the substations, such as line ID, line name, length, capacity, and status.

Line ID	Line Name	Length (miles)	Capacity	Status
101	Line 1	50	200	Active
102	Line 2	100	300	Inactive
103	Line 3	75	250	Active

3. Substations table:

This table stores information about the substations that distribute the electricity to the consumers, such as substation ID, substation name, location, capacity, and status.

Substation ID	Substation Name	Location	Capacity (MW)	Status
201	Substation 1	City 1	100	Active
202	Substation 2	City 2	200	Inactive
203	Substation 3	City 3	150	Active

4. Consumers table:

This table stores information about the consumers that use electricity from the power grid system, such as consumer ID, consumer name, location, and usage.

Consumer ID	Consumer Name	Location	Usage (MW)
301	Consumer 1	City 1	50
302	Consumer 2	City 2	100
303	Consumer 3	City 3	75

Text-to-Code (T2C) use cases

ORGANIZER

ISGF
India Smart Grid Forum

Query - What is the total capacity of active power plants that are connected to active substations in the power grid system?

SQL Query:

vbnet

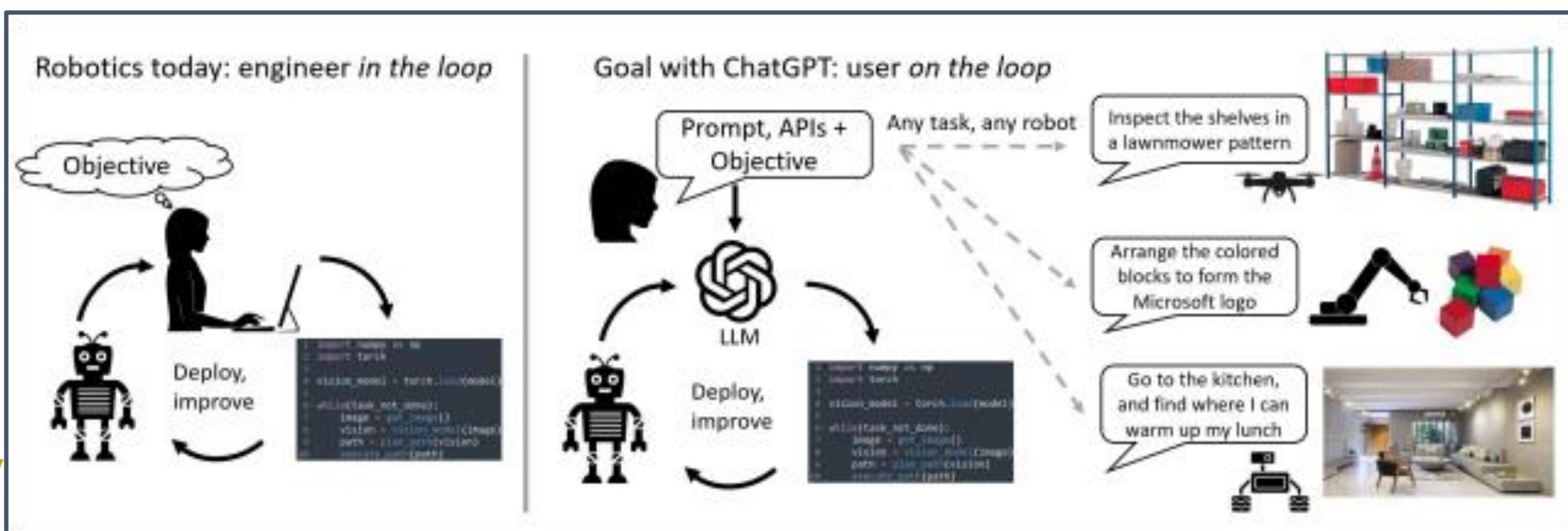
 Copy code

```
SELECT SUM(p.Capacity) AS Total_Capacity
FROM Power_Plants p
JOIN Transmission_Lines tl ON p.Plant_ID = tl.Plant_ID
JOIN Substations s ON tl.Substation_ID = s.Substation_ID
WHERE p.Status = 'Active' AND s.Status = 'Active';
```

ChatGPT for Robotics

ORGANIZER

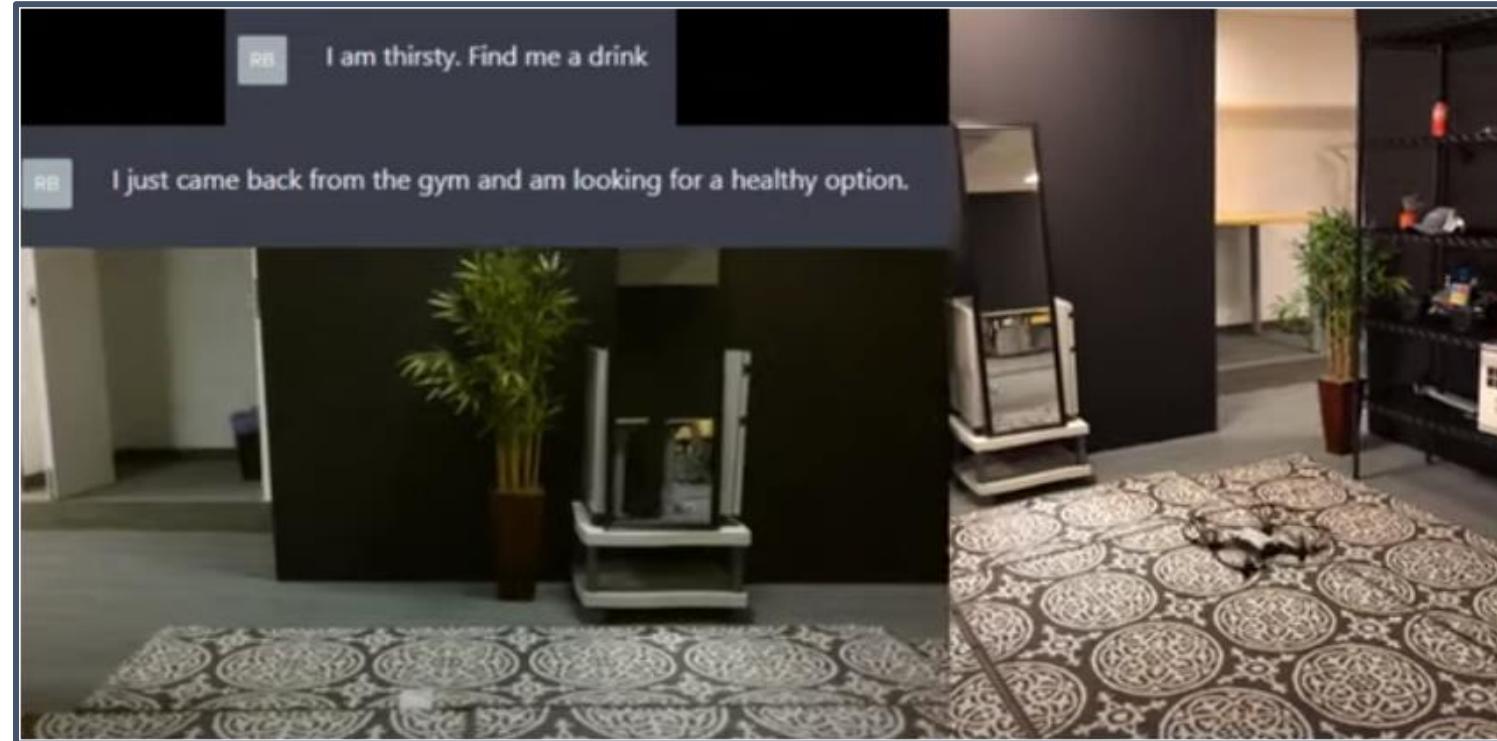
ISGF
India Smart Grid Forum



ChatGPT with Real Drone

ORGANIZER

ISGF
India Smart Grid Forum



ChatGPT with AirSim

ORGANIZER

ISGF
India Smart Grid Forum

ChatGPT with AirSim



VHDL Query

ORGANIZER

ISGF
India Smart Grid Forum

Write a VHDL code for a dynamic voltage and frequency scaling (DVFS) circuit that **adjusts the voltage and clock frequency of the design based on the processing requirements to reduce power consumption.**

ChatGPT Output

ORGANIZER

ISGF
India Smart Grid Forum

```
library ieee;
use ieee.std_logic_1164.all;
use ieee.numeric_std.all;

entity dvfs_circuit is
    port(
        clk      : in std_logic;
        rst      : in std_logic;
        load    : in std_logic_vector(15 downto 0);
        voltage : out std_logic_vector(3 downto 0);
        frequency : out std_logic_vector(7 downto 0)
    );
end entity;

architecture dvfs_arch of dvfs_circuit is

    constant max_load    : integer := 65535;
    constant min_voltage : integer := 0;
    constant max_voltage : integer := 15;
    constant max_freq    : integer := 255;
    constant min_freq    : integer := 10;
    constant alpha        : integer := 1; -- voltage step
    constant beta        : integer := 10; -- frequency step

    signal current_voltage  : integer range min_voltage to max_voltage := 0;
    signal current_frequency : integer range min_freq to max_freq := 0;
```

ChatGPT Output

ORGANIZER

ISGF
India Smart Grid Forum

```
begin

    -- the input load signal is normalized to 0-1 range
    signal norm_load : integer range 0 to 100 := 0;
    norm_load <= integer(round(unsigned(load) * 100 / max_load));

    -- voltage and frequency update process
    process(clk, rst)
    begin
        if rst = '1' then
            current_voltage <= 0;
            current_frequency <= 0;
            voltage <= std_logic_vector(to_unsigned(current_voltage, voltage'length));
            frequency <= std_logic_vector(to_unsigned(current_frequency, frequency'length));
        elsif rising_edge(clk) then
            if norm_load < 30 then -- low processing requirements
                current_voltage <= current_voltage - alpha;
                current_frequency <= current_frequency - beta;
            elsif norm_load > 70 then -- high processing requirements
                current_voltage <= current_voltage + alpha;
                current_frequency <= current_frequency + beta;
            end if;
        end if;
    end process;
end architecture;
```

ChatGPT Output

ORGANIZER



```
-- voltage and frequency limits
if current_voltage < min_voltage then
    current_voltage <= min_voltage;
elsif current_voltage > max_voltage then
    current_voltage <= max_voltage;
end if;

if current_frequency < min_freq then
    current_frequency <= min_freq;
elsif current_frequency > max_freq then
    current_frequency <= max_freq;
end if;

voltage <= std_logic_vector(to_unsigned(current_voltage, voltage'length));
frequency <= std_logic_vector(to_unsigned(current_frequency, frequency'length));
end if;
end process;

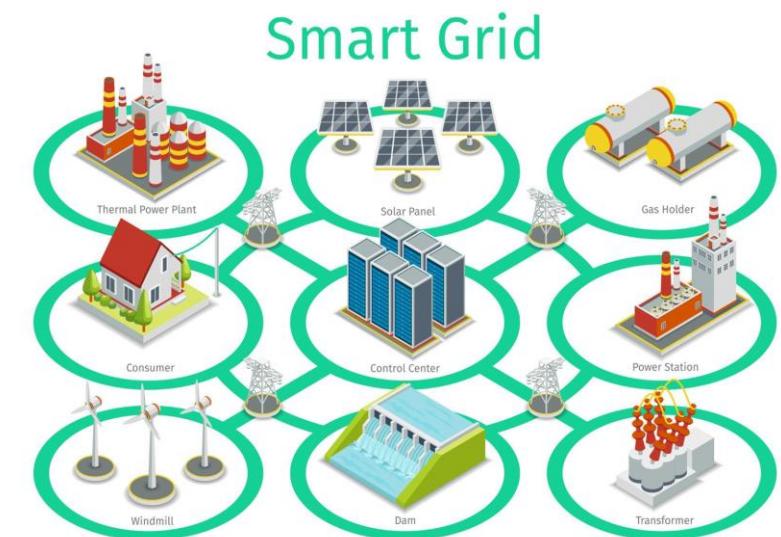
end architecture;
```

Smart Grid Problem

ORGANIZER

ISGF
India Smart Grid Forum

Suppose you are a smart grid operator responsible for managing the energy distribution in a city. You have three energy sources available: solar panels, wind turbines, and a power plant. The power plant has a fixed cost of \$10,000 per day to operate, while the solar panels and wind turbines have variable costs based on their output. The cost of generating energy from solar panels is \$50 per unit, while the cost of generating energy from wind turbines is \$75 per unit. Your goal is to meet a daily energy demand of 5,000 units at the lowest possible cost.



ChatGPT Response

ORGANIZER

ISGF
India Smart Grid Forum

```
import pulp

def optimize_smart_grid(solar_cost, wind_cost, power_plant_cost, solar_max, wind_m
    # Define the problem as a linear programming (LP) problem
    problem = pulp.LpProblem("Smart Grid Optimization", pulp.LpMinimize)

    # Define decision variables
    solar_energy = pulp.LpVariable("Solar Energy", lowBound=0, cat="Continuous")
    wind_energy = pulp.LpVariable("Wind Energy", lowBound=0, cat="Continuous")
    power_plant_energy = pulp.LpVariable("Power Plant Energy", lowBound=0, cat="Co

    # Define objective function
    problem += solar_cost * solar_energy + wind_cost * wind_energy + power_plant_c

    # Define constraints
    problem += solar_energy + wind_energy + power_plant_energy == demand
    problem += solar_energy <= solar_max
    problem += wind_energy <= wind_max

    # Solve the LP problem
    problem.solve()

    # Print the results
    print("Optimal energy generation plan:")
    print("Solar Energy: ", solar_energy.value(), "units")
    print("Wind Energy: ", wind_energy.value(), "units")
    print("Power Plant Energy: ", power_plant_energy.value(), "units")
    print("Total cost: $", pulp.value(problem.objective))
```

Audio-to-Text (A2T)

ORGANIZER

ISGF
India Smart Grid Forum

Audio-to-text generative AI, also known as speech-to-text, is the process of converting spoken language in an audio recording into written text. This technology uses artificial intelligence models to transcribe the audio and produce a textual representation of the speech.



Audio-to-Text (A2T) Use cases

ORGANIZER

ISGF
India Smart Grid Forum

A2T Core Capabilities



Transcription



Closed
captioning



Voice search



Call center
analytics



Language
Learning

A2T Use Cases for Energy Sector



Maintenance and
repair
documentation



Safety incident
reporting



Training and
knowledge
management



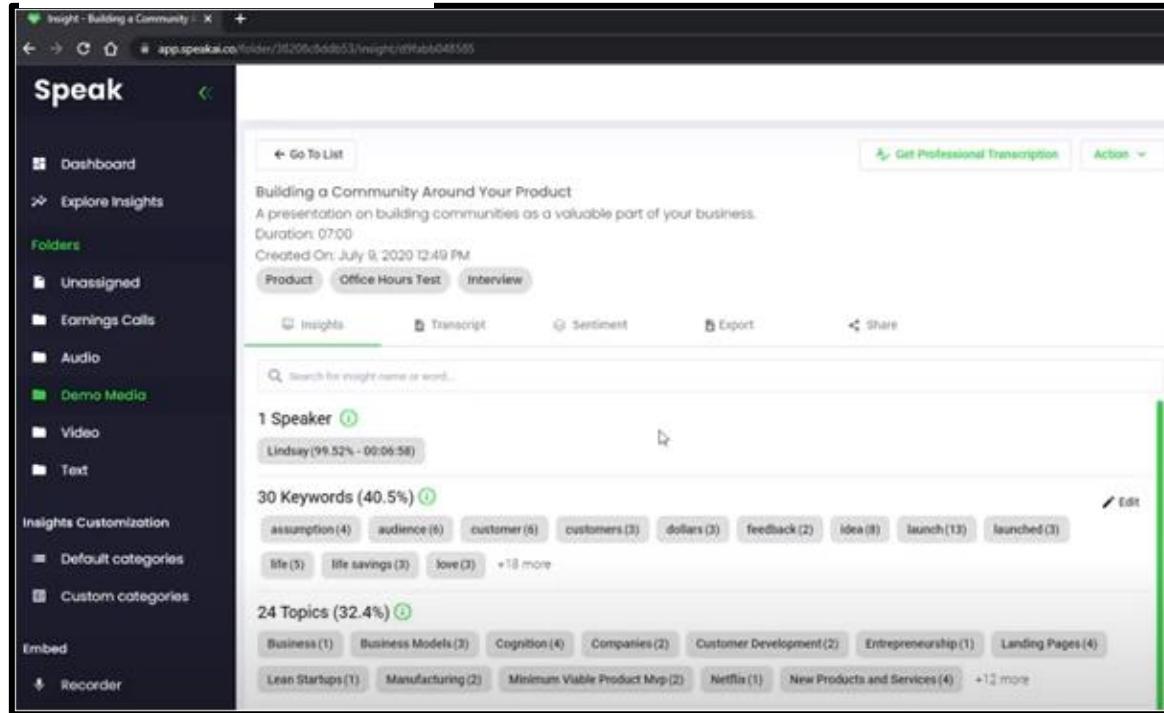
Regulatory
compliance



Customer Support
Voicebots

Audio-to-Text (A2T) Tools

Speak

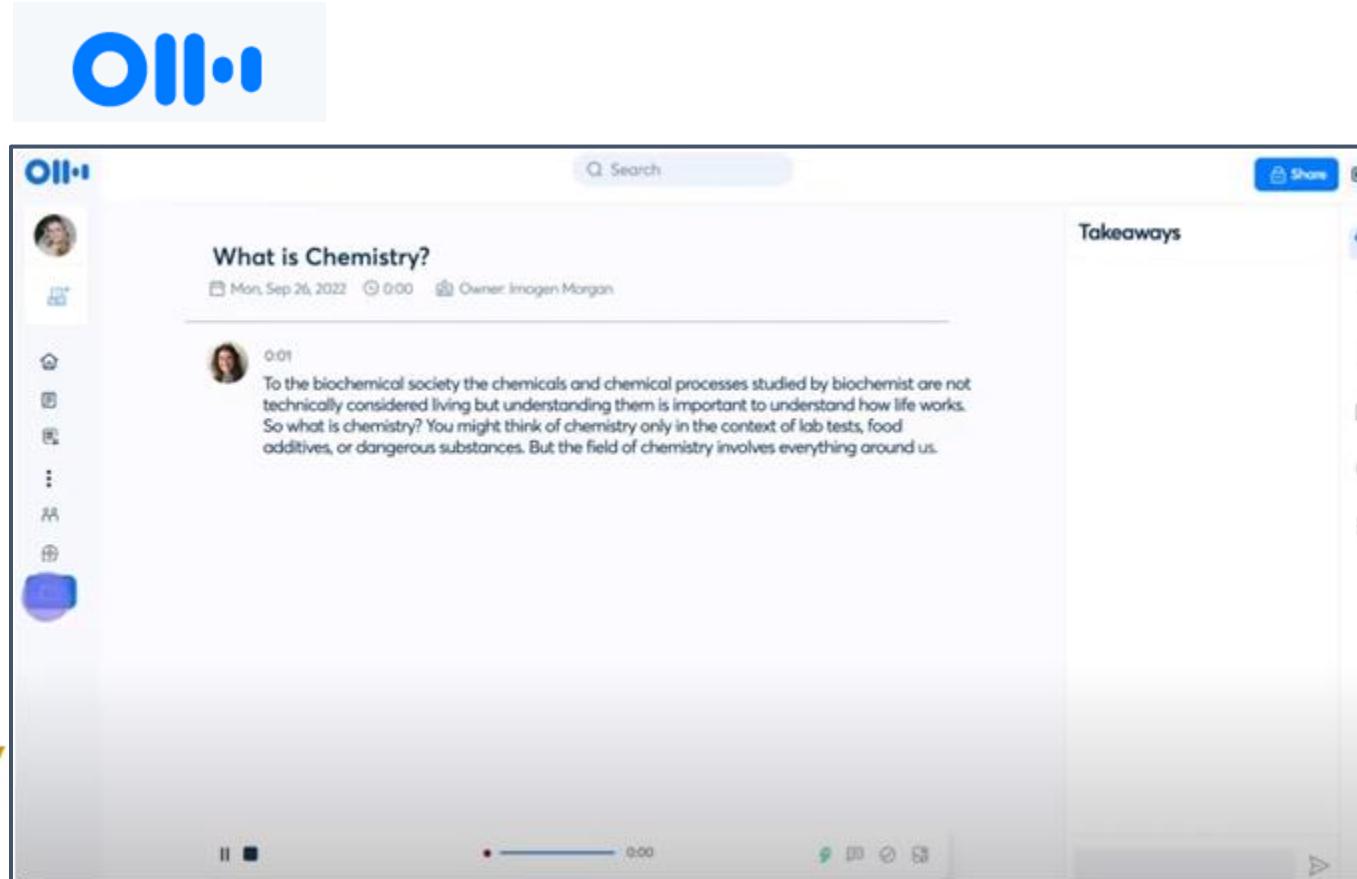


The screenshot shows the 'Insight - Building a Community' page on the Speak AI platform. The left sidebar includes sections for Dashboard, Explore Insights, Folders (Unassigned, Earnings Calls, Audio, Demo Media, Video, Text), Insights Customization (Default categories, Custom categories), and Embed. The main content area displays an insight titled 'Building a Community Around Your Product' with a duration of 07:00, created on July 9, 2020. It features tabs for Product, Office Hours Test, and Interview, and buttons for Go To List, Get Professional Transcription, Action, Insights, Transcript, Sentiment, Export, and Share. Below the tabs, there's a search bar and a speaker icon. The analysis results show 1 Speaker (Lindsay), 30 Keywords (40.5%), and 24 Topics (32.4%). The keywords listed include assumption(4), audience(6), customer(8), customers(3), dollars(3), feedback(2), idea(8), launch(13), launched(3), life(5), life savings(3), love(3), +18 more. The topics listed include Business(1), Business Models(3), Cognition(4), Companies(2), Customer Development(2), Entrepreneurship(1), Landing Pages(4), Lean Startups(1), Manufacturing(2), Minimum Viable Product Mvp(2), Netflix(1), New Products and Services(4), +12 more.

Main Features:

- Named entity recognition
- Deep search
- APIs and integrations
- Media management
- Dashboard reports and audio capture

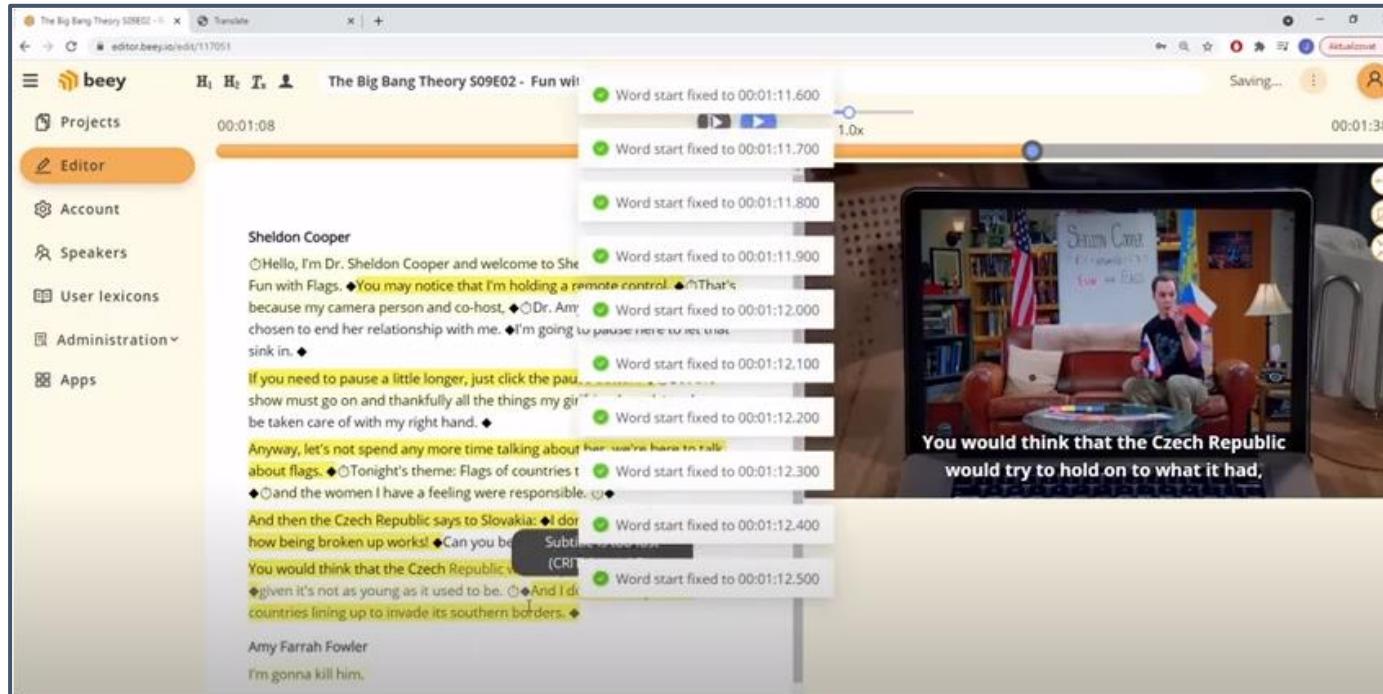
Audio-to-Text (A2T) Tools



The screenshot shows the Otter.ai application interface. At the top, there's a search bar and a 'Show' button. On the left, a sidebar includes a user profile picture, a 'Takeaways' section with a diamond icon, and various other icons. The main content area displays a transcription titled 'What is Chemistry?' from Monday, September 26, 2022, at 0:00, owned by Imogen Morgan. The transcription text reads: 'To the biochemical society the chemicals and chemical processes studied by biochemist are not technically considered living but understanding them is important to understand how life works. So what is chemistry? You might think of chemistry only in the context of lab tests, food additives, or dangerous substances. But the field of chemistry involves everything around us.' Below the transcription, there are playback controls (play/pause, volume, seek bar at 0:00), and a 'Takeaways' section on the right side.

Main Features:

- Intuitive and well-designed
- Available on desktop and mobile
- Manage directly in-app
- Audio playback at different speeds
- Automatically transcribe conversations

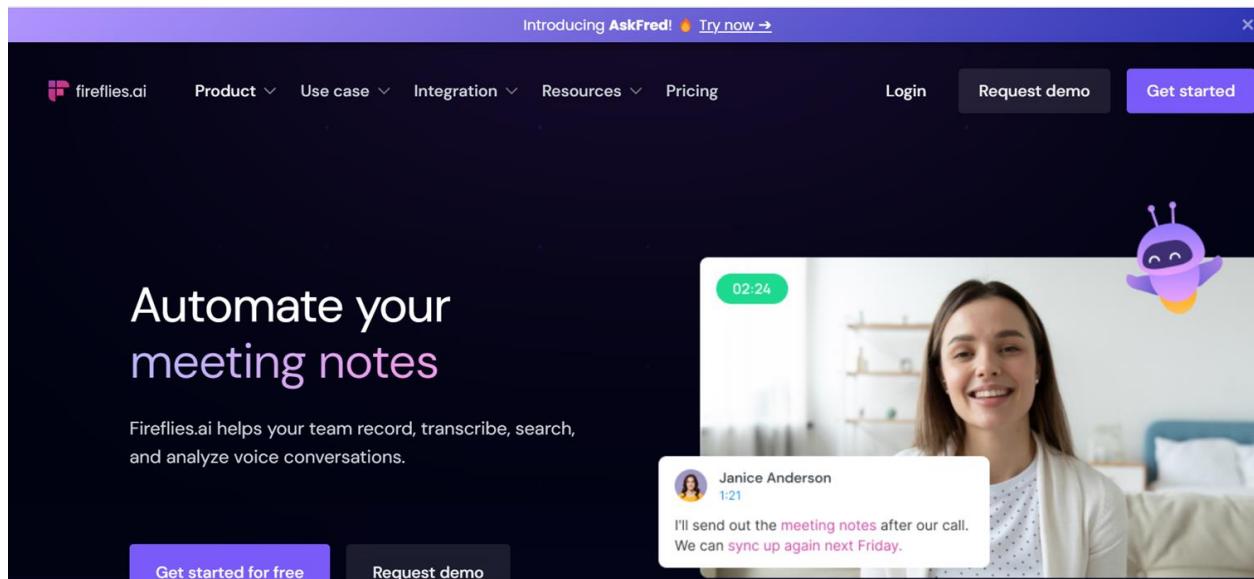
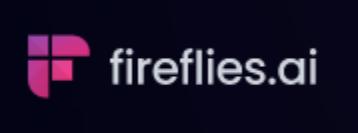


Audio-to-Text (A2T) Tools

Main Features:

- Intuitive and well-designed
- Lightning fast execution
- Allows manual editing to correct errors
- Supports 20 Languages

Audio-to-Text (A2T) Tools



Introducing AskFred! Try now →

fireflies.ai Product Use case Integration Resources Pricing Login Request demo Get started

Automate your meeting notes

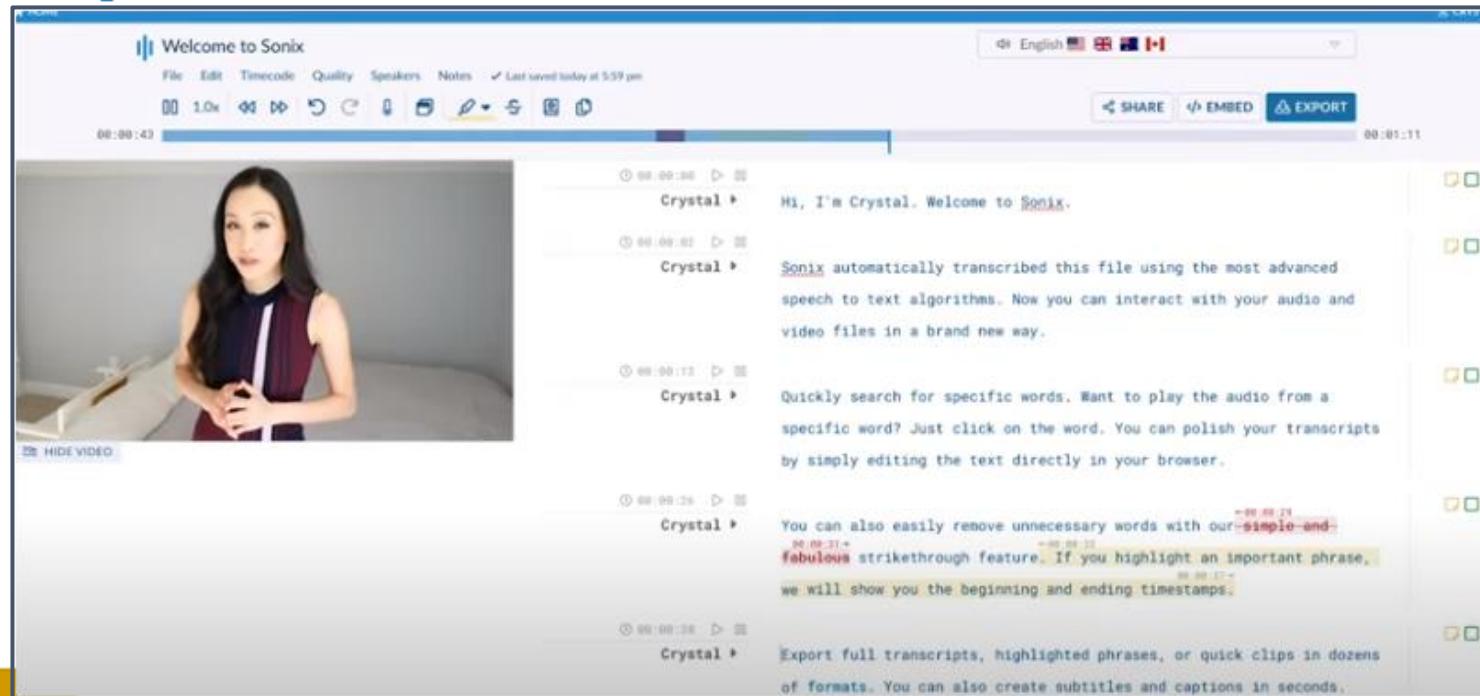
Fireflies.ai helps your team record, transcribe, search, and analyze voice conversations.

Get started for free Request demo

Main Features:

- Meeting bot that can auto join calls
- Chrome extension
- Transcribe existing audio files inside the dashboard
- Instantly record meetings
- Skim transcripts while listening to audio

Audio-to-Text (A2T) Tools



The screenshot shows the Sonix software interface. On the left, there is a video player window displaying a woman with long dark hair, wearing a purple top, standing in front of a white wall. On the right, the transcript area shows several lines of text with corresponding timestamps and speaker labels. The transcript includes:

- 00:00:43 Crystal > Hi, I'm Crystal. Welcome to Sonix.
- 00:00:52 Crystal > Sonix automatically transcribed this file using the most advanced speech to text algorithms. Now you can interact with your audio and video files in a brand new way.
- 00:00:57 Crystal > Quickly search for specific words. Want to play the audio from a specific word? Just click on the word. You can polish your transcripts by simply editing the text directly in your browser.
- 00:00:58 Crystal > You can also easily remove unnecessary words with our ~~simple-end~~ ~~Fabulous~~ strikethrough feature. If you highlight an important phrase, we will show you the beginning and ending timestamps.
- 00:00:59 Crystal > Export full transcripts, highlighted phrases, or quick clips in dozens of formats. You can also create subtitles and captions in seconds.

Main Features:

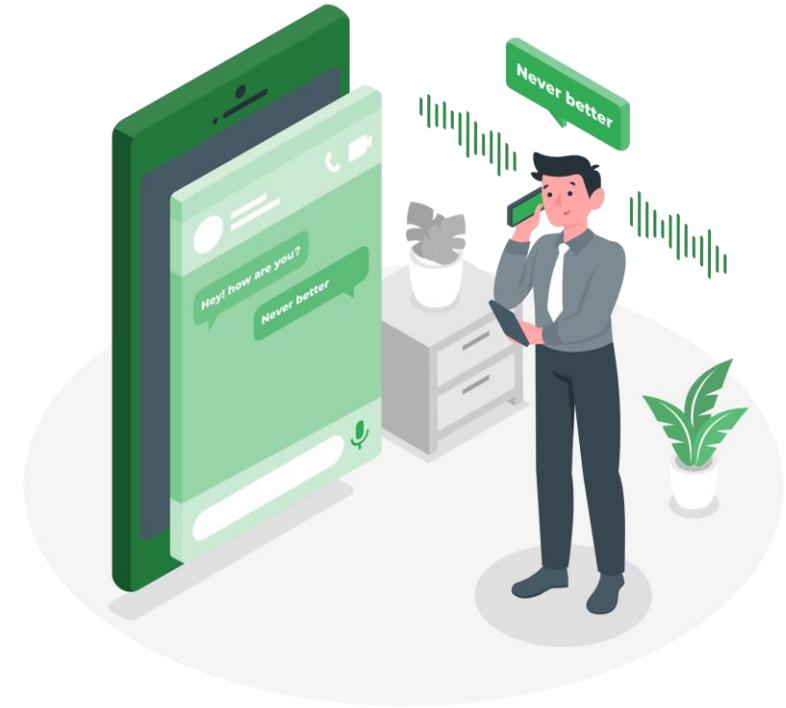
- Highlights words and identifies accuracy confidence
- Multi-user capability
- Transcribes 30 minutes of audio in 3-4 minutes
- Drag and drop
- Speaker labeling

Text-to-Audio (T2A)

ORGANIZER

ISGF
India Smart Grid Forum

Text-to-Audio (T2A) AI refers to an artificial intelligence system that uses text-to-speech (TTS) technology to convert written text into spoken audio. T2A AI systems use advanced algorithms and machine learning techniques to analyze written text and generate an artificial voice that reads out the text in a natural and human-like manner.



Text-to-Audio (T2A)

ORGANIZER

ISGF
India Smart Grid Forum

Audio captioning

Generating captions to describe audio content

E-book reading

Reading e-books out loud

Voice synthesis

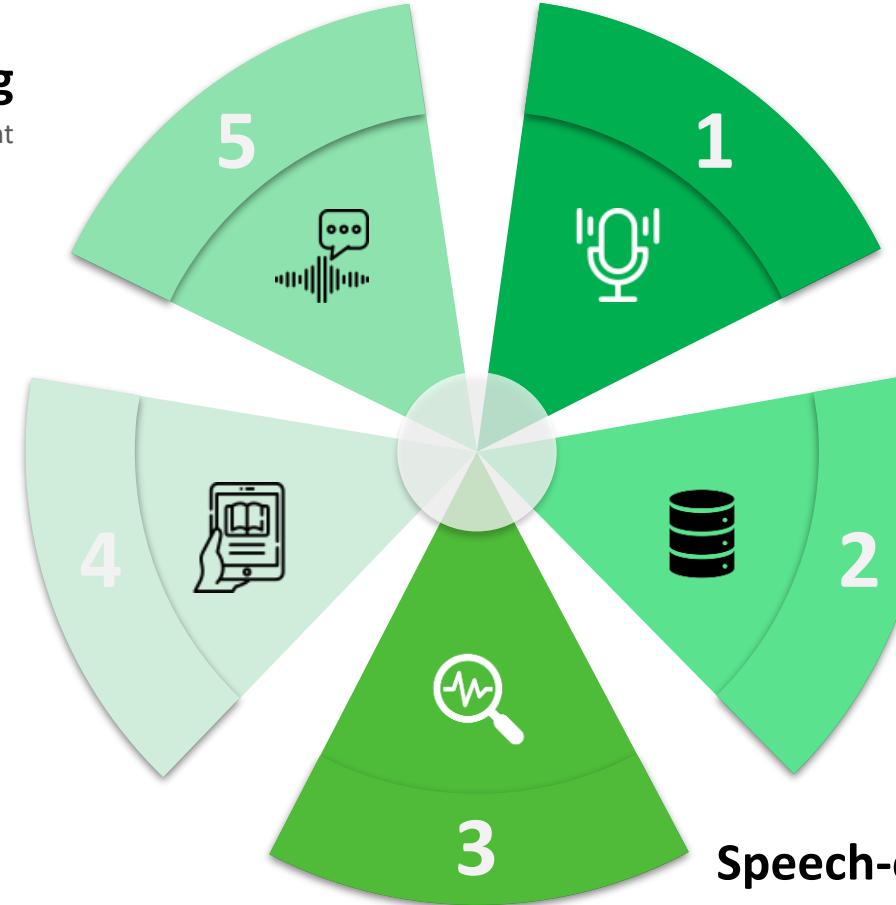
Generating a human-like voice from written text

Voice-enabled applications

Creating applications that respond to voice commands

Speech-enabled devices

Enabling devices to respond to voice commands and provide spoken feedback



Text-to-Audio (T2A) Use cases

ORGANIZER

ISGF
India Smart Grid Forum

T2A Core Capabilities

- » Accessibility
- » Language Learning
- » Personal Assistants
- » Audiobook Production
- » Customer Service
- » Navigation Systems

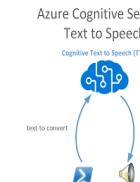
T2A Use Cases for Energy Sector

- » Safety and emergency communications
- » Asset management and monitoring
- » Control room operations
- » Regulatory compliance
- » Maintenance and repair instructions
- » Training and knowledge management

Text-to-Audio (T2A) Tools

ORGANIZER

ISGF
India Smart Grid Forum



Notevibes



Amazon Polly



 ispring suite

 isuw@isuw.in

 www.isuw.in

 @ISUW_India

 India Smart Utility Week (ISUW)

Voice Synthesis



podcast.ai



 Listnr

REPLICA

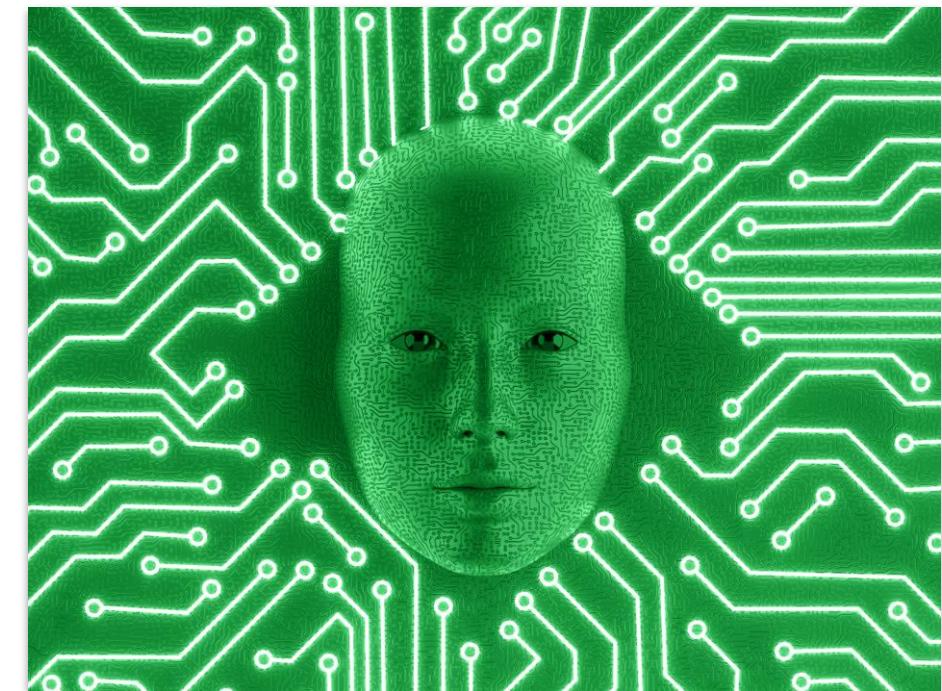


Text-to-Image (T2I)

ORGANIZER

ISGF
India Smart Grid Forum

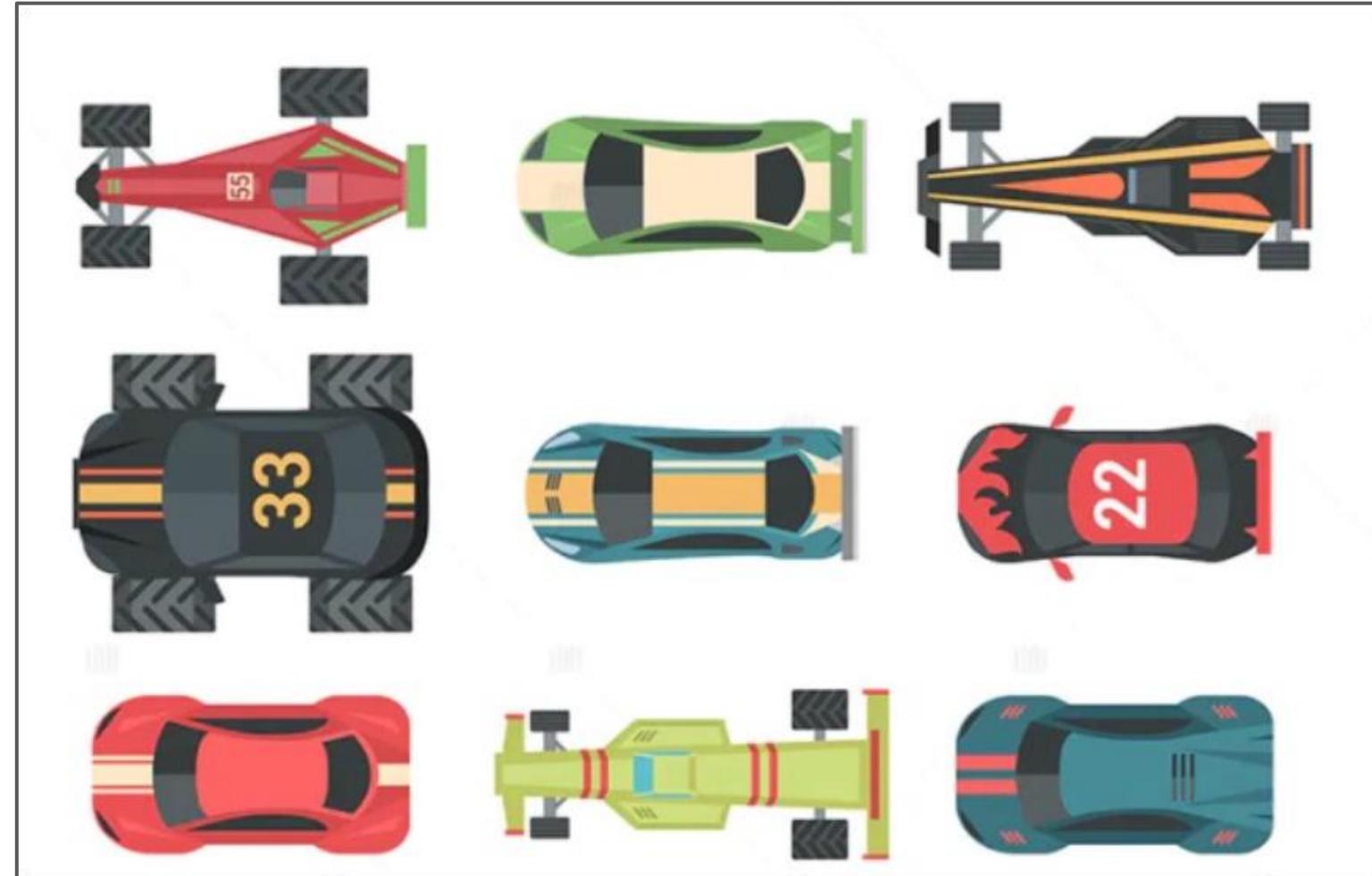
Text-to-Image Generative AI refers to a type of artificial intelligence that uses deep learning techniques to generate images from textual descriptions. The technology is based on the idea that a machine can learn to generate images by analyzing large datasets of images and their corresponding textual descriptions.

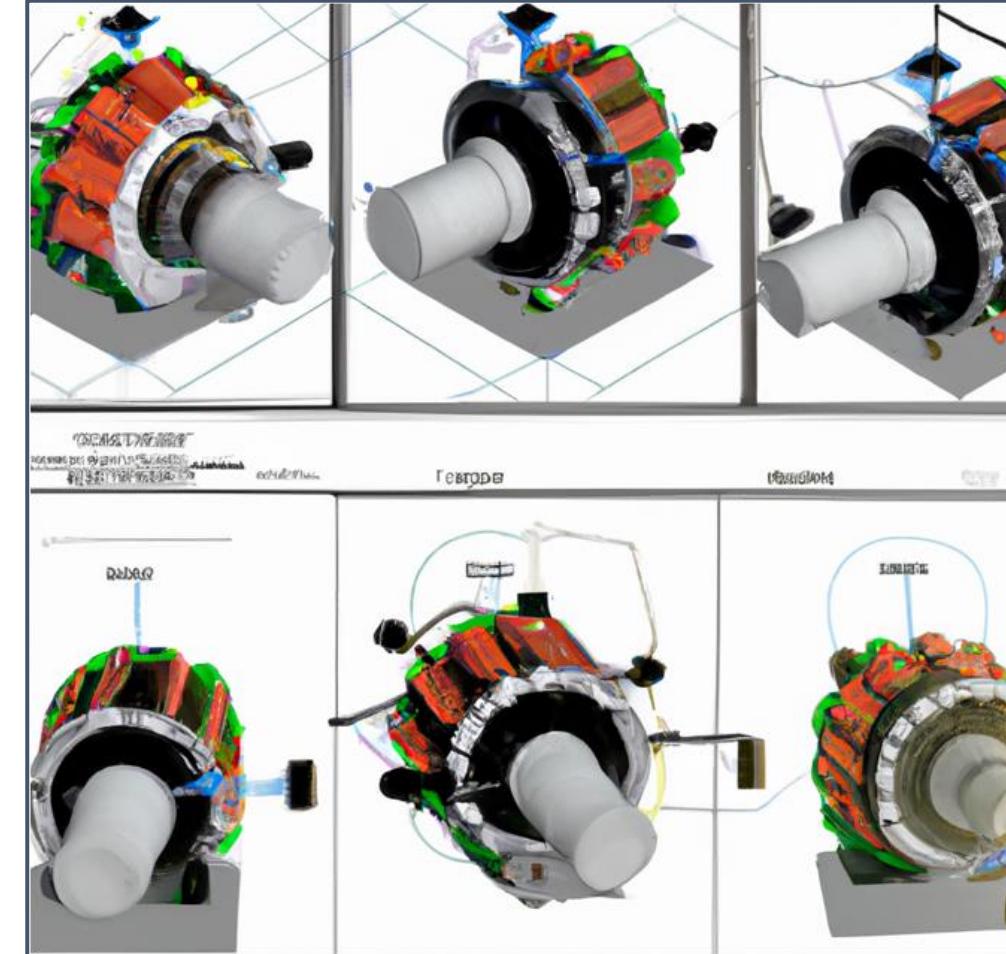


Text-to-Image (T2I): Translation

ORGANIZER

ISGF
India Smart Grid Forum







India
SMART UTILITY
Week 2023

Text-to-Image (T2I): Envisioning Future

ORGANIZER

ISGF
India Smart Grid Forum

28 Feb – 04 March 2023 | New Delhi



isuw@isuw.in

www.isuw.in

@ISUW_India India Smart Utility Week (ISUW)



India
SMART UTILITY
Week 2023

Text-to-Image (T2I): Envisioning Future

ORGANIZER

ISGF
India Smart Grid Forum

28 Feb – 04 March 2023 | New Delhi



Text-to-Image (T2I) Tools



Text-to-Image (T2I) Tools

ORGANIZER

ISGF
India Smart Grid Forum

Design

 **Diagram**

DeepAI

VIZCOM

 Google Colaboratory

 **uizard**

 **Poly**

 **AISEO**

NightCafe

 **maket**

CALA

INTERIORA

Stable Diffusion

 **Hypotenuse AI**

artbreeder

 **OpenAI**

Media/Advertising

SALT

 **THE CULTURE DAO**

 **fotor**

Consumer/Social

 **MidJourney**

 **DEEP DREAM GENERATOR**



Anonymizer

Image Generation

 **MidJourney**

 **Open Art**

 **craiyon**

 **PLAYGROUND**

 **WOMBORAI**

 **PhotoRoom**

 **ROSEBUD.AI**

 **alpaca**

 **Lexica**

 **Nyx gallery**

 **artbreeder**

 **starryai.**

Image-to-Text (I2T)

ORGANIZER

ISGF
India Smart Grid Forum

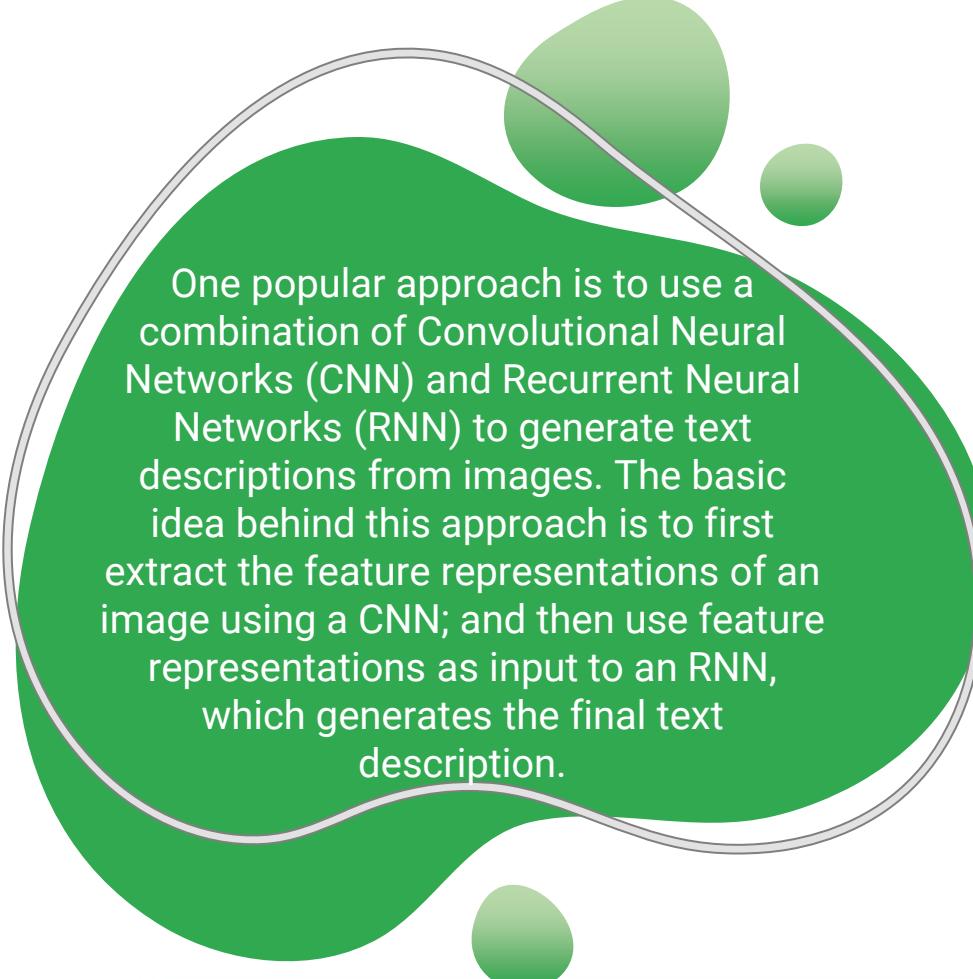
Image to text Generative AI is a type of artificial intelligence that uses deep learning models to generate text descriptions of images. This technology is also known as image captioning, and it works by analyzing the visual features of an image and generating a natural language description of those features.



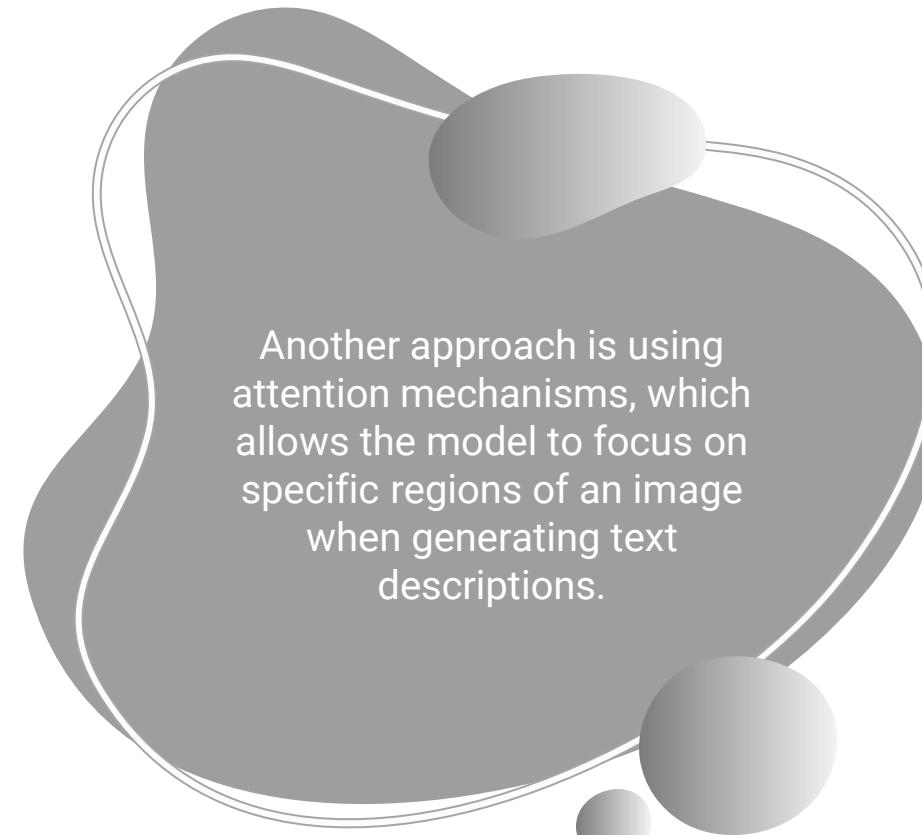
Image-to-Text (I2T)

ORGANIZER

ISGF
India Smart Grid Forum



One popular approach is to use a combination of Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) to generate text descriptions from images. The basic idea behind this approach is to first extract the feature representations of an image using a CNN; and then use feature representations as input to an RNN, which generates the final text description.



Another approach is using attention mechanisms, which allows the model to focus on specific regions of an image when generating text descriptions.

Image-to-Text (I2T) Tools

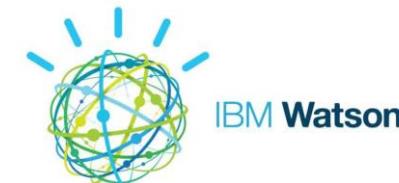
ORGANIZER

ISGF
India Smart Grid Forum

Image Captioning



Brandwatch



I2T Core Capabilities

- Using neural networks to generate text from images.
- Training data models to use attention mechanism to describe specific areas of images with text.

I2T Use Cases for Energy Sector

- Text inputs from images for predictive maintenance
- Strategic energy planning and decision making based on text insights generated from images. Reporting and analysis inputs

Text-to-Video (T2V)

ORGANIZER

ISGF
India Smart Grid Forum

Text to video generative AI is a type of artificial intelligence that is able to generate video content from text input. In this context, the text input can be a script, a description of a scene, or any other type of textual input that describes the desired video output. The AI system then generates a video that matches the input text as closely as possible.

Text-to-Video (T2V)

ORGANIZER

ISGF
India Smart Grid Forum



One popular approach is to use a combination of deep learning and computer vision techniques to generate videos from text descriptions.



One way to approach T2V is to first generate an image from the text description using a T2I model, and then use that image as the keyframe for the video. Additional frames can then be generated by applying motion and other visual transformations to the keyframe.



Another approach is using Generative Adversarial Networks (GANs) to generate videos from text descriptions. GANs consist of two neural networks - a generator network and a discriminator network. The generator network generates videos from text descriptions, while the discriminator network tries to distinguish the generated videos from real videos. The two networks are trained together to improve the quality of the generated videos.

Text-to-Video (T2V) use cases

ORGANIZER

ISGF
India Smart Grid Forum

T2V Core Capabilities

- Generate videos from text description by using deep learning and computer vision.
- Using GAN neural networks to generate video from text.
- Fine-tuning language models like GPT-Video & Dalle-4 for generating video from text.

T2V Use Cases for Energy Sector

- Explaining complex processes
- Creating safety training videos
- Advertising and marketing
- Reporting and analysis
- Visualization of renewable energy projects

Text-to-Video (T2V) Tools



Editing/Generation



Opus



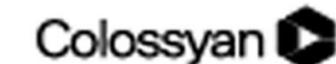
Synths Video



Personalized
Videos



Hour One.



Text-to-Motion (T2M)

ORGANIZER

ISGF
India Smart Grid Forum

Text-to-Motion (T2M) is an NLP task that involves generating animation or motion from a given text description.

This task can be used to generate animations of objects, characters, or even abstract concepts that are described in natural language.



One approach to T2M is to use a combination of deep learning and computer animation techniques to generate animations from text descriptions.



A T2M generation model can be trained to understand the text input and generate a corresponding animation.



A related approach is to use motion capture data, where an actor's movements are recorded and then used to animate a digital character. The text input can be used to guide the selection of the appropriate motion capture data.



Another approach is to use a physics engine to simulate the motion of objects in a scene. The text input can be used to specify the initial conditions and constraints of the simulation. There are also several pre-trained models like DALL-E 4, which can be fine-tuned for T2M tasks.

Text-to-Motion (T2M) Tools

ORGANIZER

ISGF
India Smart Grid Forum

1. MDM: Human Motion Diffusion Model

1. T2M-GPT

1. HumanML3D/Text2motion

Text-to-Motion (T2M) Tools

ORGANIZER

ISGF
India Smart Grid Forum

MDM: Human Motion Diffusion Model

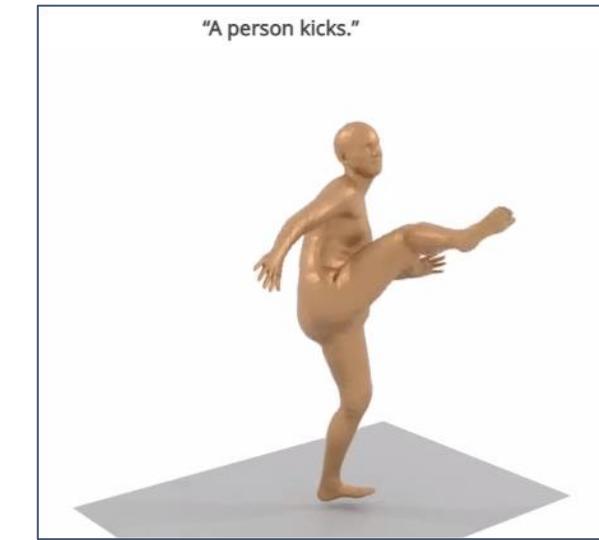
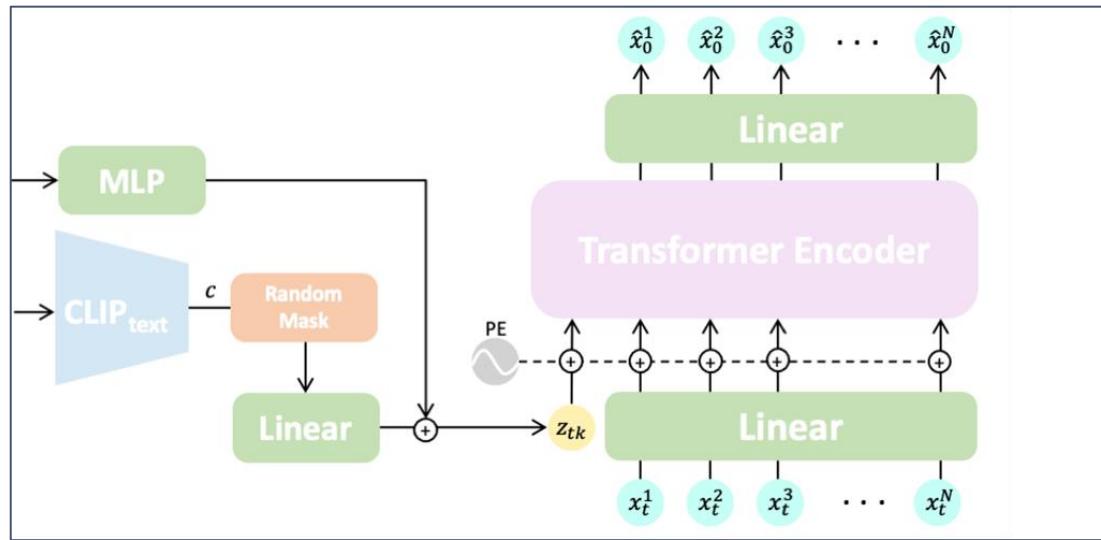
MDM is transformer-based, combining insights from motion generation literature. A notable design-choice is the prediction of the sample, rather than the noise, in each diffusion step. This facilitates the use of established geometric losses on the locations and velocities of the motion, such as the foot contact loss. MDM is a generic approach, enabling different modes of conditioning, and different generation tasks. Model is trained with lightweight resources and yet achieves state-of-the-art results on leading benchmarks for text-to-motion and action-to-motion.

Text-to-Motion (T2M) Tools

ORGANIZER

ISGF
India Smart Grid Forum

MDM: Human Motion Diffusion Model



Text-to-Motion (T2M) Tools

ORGANIZER

ISGF
India Smart Grid Forum

T2M-GPT

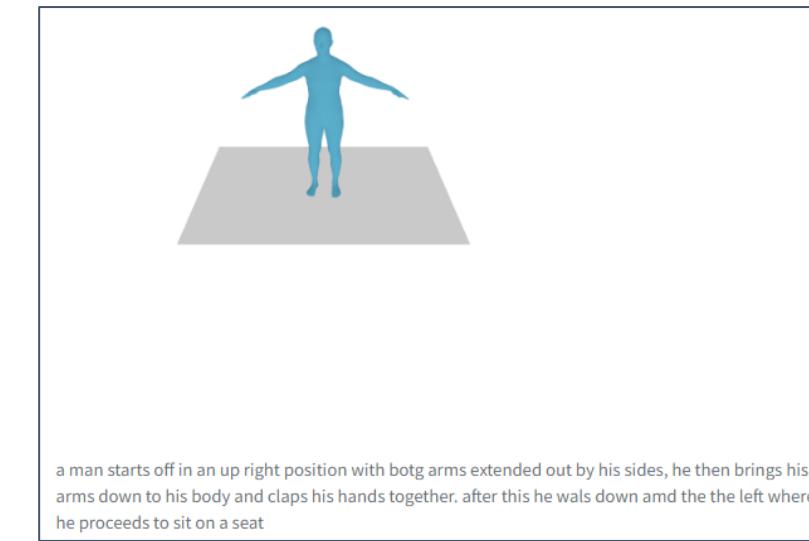
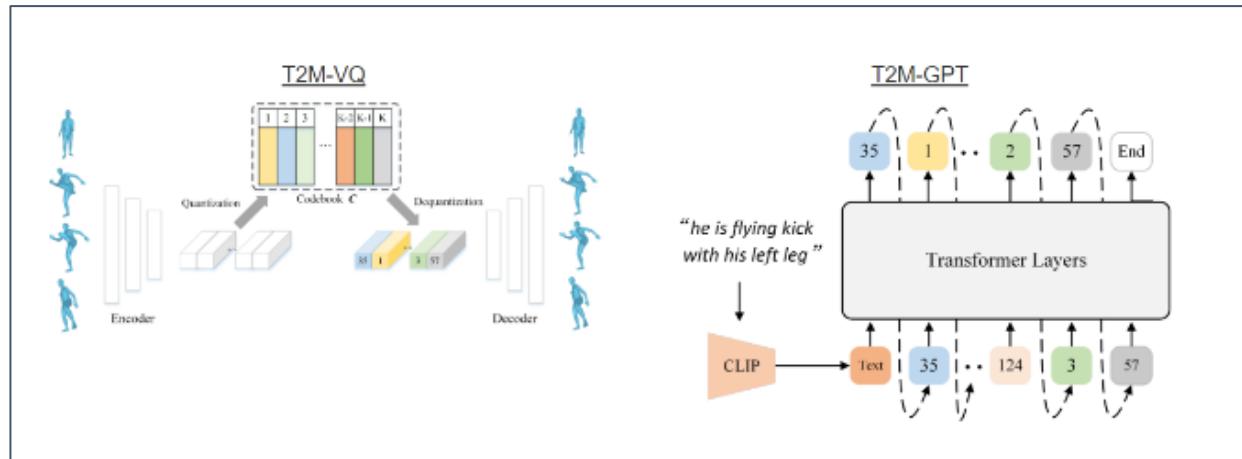
T2M-GPT is a conditional generative framework based on Vector Quantised-Variational AutoEncoder (VQ-VAE) and Generative Pre-trained Transformer (GPT) for human motion generation from textural descriptions. A simple CNN-based VQ-VAE with commonly used training recipes (EMA and Code Reset) allows to obtain high-quality discrete representations. T2M-GPT shows better performance than competitive approaches, including recent diffusion-based approaches.

Text-to-Motion (T2M) Tools

ORGANIZER

ISGF
India Smart Grid Forum

T2M-GPT



Text-to-Motion (T2M) Tools

ORGANIZER

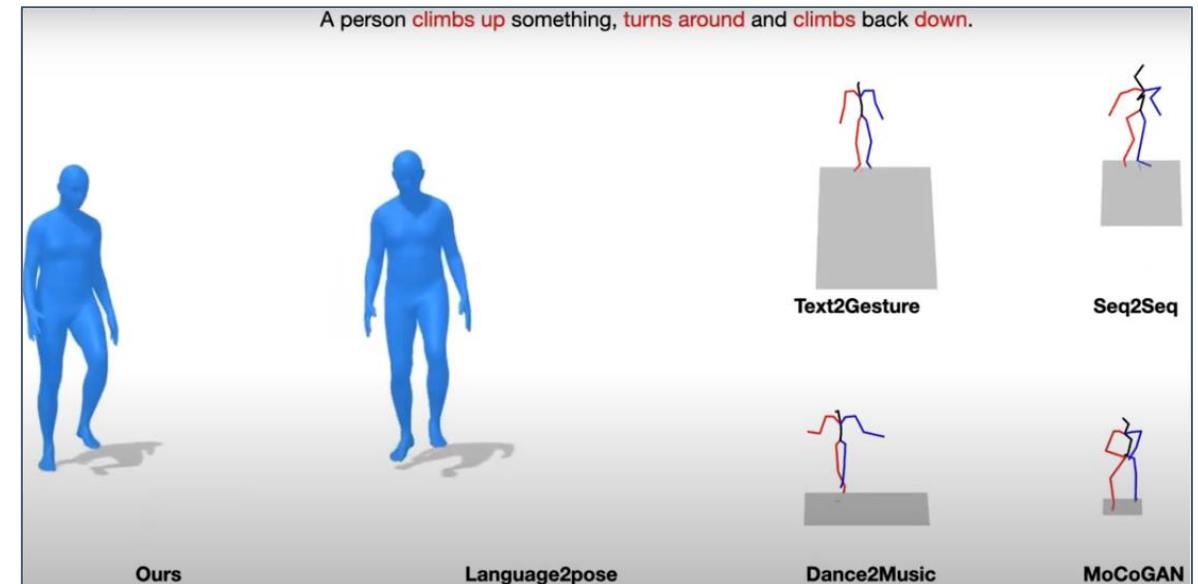
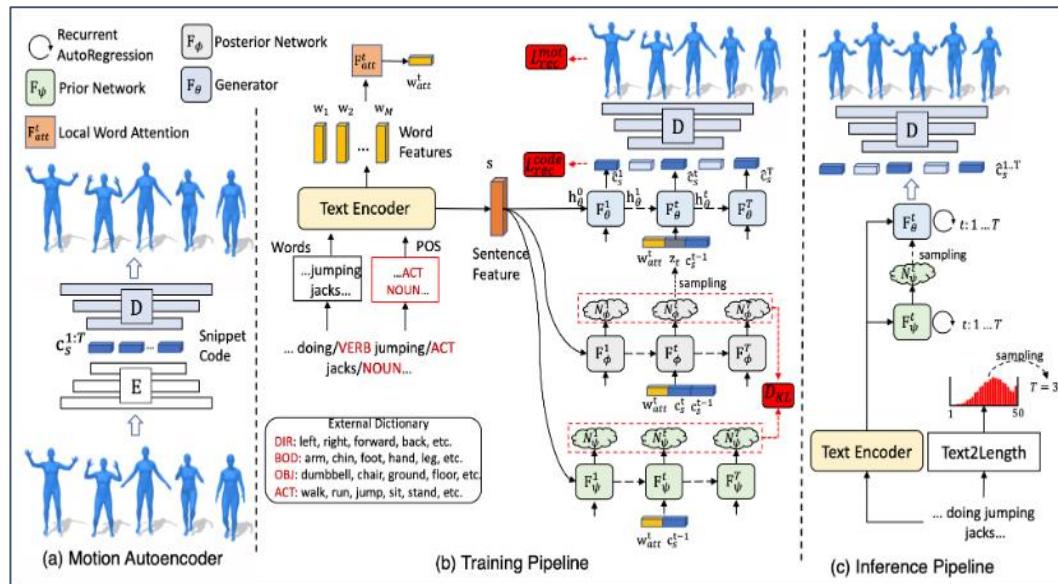
ISGF
India Smart Grid Forum

HumanML3D/Text2motion

The learned distribution function of motion lengths conditioned on the input text is sampled by Text2length. This is followed by the use of a temporal variational autoencoder by the text2motion module to synthesize a diverse set of human motions of the sampled lengths. Motion snippet code is proposed as the internal motion representation, which captures local semantic motion contexts and is empirically shown to facilitate the generation of plausible motions faithful to the input text, instead of directly engaging with pose sequences. Additionally, a large-scale dataset of scripted 3D human motions, HumanML3D, consisting of 14,616 motion clips and 44,970 text descriptions, has been constructed. The effectiveness of the approach is demonstrated by extensive empirical experiments.

Text-to-Motion (T2M) Tools

HumanML3D/Text2motion



Brain-to-Text (B2T)

ORGANIZER

ISGF
India Smart Grid Forum

Brain to text generative AI is a type of artificial intelligence that aims to decode human brain activity into text, bypassing the need for actual speech or typing. The technology works by analyzing patterns of brain activity using brain-computer interface (BCI) devices, and then generating corresponding text output based on those patterns.

Brain-to-Text (B2T)

ORGANIZER

ISGF
India Smart Grid Forum

- One approach to B2T is to use electroencephalography (EEG) to record brain activity, and then use machine learning algorithms to analyze the EEG data and transcribe it into text. This approach has been used in research to transcribe simple phrases, such as "left hand" or "right hand".
- Another approach is to use functional magnetic resonance imaging (fMRI) to record brain activity, and then use machine learning algorithms to analyze the fMRI data and transcribe it into text. This approach has been used in research to transcribe short phrases or sentences.
- It is important to note that the technology for B2T is still in its early stages and there is a lot of ongoing research in this field. The transcription of thoughts into text is still a difficult task and requires a lot of research and development.

Brain-to-Text (B2T) Tools

ORGANIZER

BrainGate: BrainGate is a research collaboration that is developing brain-computer interfaces to help people with paralysis or other physical disabilities to communicate and interact with the world around them. The system uses a small implant that is placed in the brain to record neural activity, which is then analyzed by a computer to generate text or other outputs.

Neuralink: Neuralink is a company founded by Elon Musk that is developing brain-machine interfaces for a variety of applications, including medical treatments and enhancing human intelligence. The company is working on a system that uses a small implant in the brain to record neural activity, which can be used to control computers or other devices.

CTRL-Labs: CTRL-Labs is a startup that is developing a wristband that can detect the electrical signals generated by the brain when a person thinks about moving their hands or fingers. The system uses machine learning algorithms to translate these signals into commands that can be used to control a computer or other device.



Brain-to-Text (B2T) Use cases

ORGANIZER

ISGF
India Smart Grid Forum

B2T Core Capabilities

- Using ML algorithms to record brain activity and ECG data into text.
- Using ML algorithms to transcribe FMRI data into text.

B2T Use Cases

- Monitoring and Diagnosing Worker Health
- Enhancing Human-Machine Interfaces
- Training and Knowledge Management

Audio-to-Audio (A2A)

ORGANIZER

ISGF
India Smart Grid Forum

Audio-to-audio AI refers to artificial intelligence systems that are designed to process and transform audio signals in various ways. These systems use machine learning algorithms to analyze audio data and generate output audio that is modified or enhanced in some way.

Audio-to-Audio (A2A)

ORGANIZER

ISGF
India Smart Grid Forum

One example of A2A is Audio Source Separation, which allows one to isolate different sounds from individual sources. For example, if we have an audio file with multiple people speaking, we can get an audio file for each of them. We can then use an Automatic Speech Recognition system to extract the text from each of these sources as an initial step for our system..

Audio-to-Audio can also be used to remove noise from audio files: we get one audio for the person speaking and another audio for the noise. This can also be useful when we have multi-person audio with some noise: we can get one audio for each person and then one audio for the noise.

Another use case is Audio Style Transfer, where we can transfer the style of one audio recording to another. For example, we can transfer the singing style of a famous singer to a different song.

Audio-to-Audio (A2A) Use cases

ORGANIZER

ISGF
India Smart Grid Forum

A2A Core Capabilities

- Speech/Audio synthesis
- Speech/Audio enhancement
- Audio style transfer.
- Speech recognition and natural language processing
- Noise reduction
- Music production and remixing
- Sound effects and audio synthesis
- Voice conversion
- Audio compression
- Audio restoration

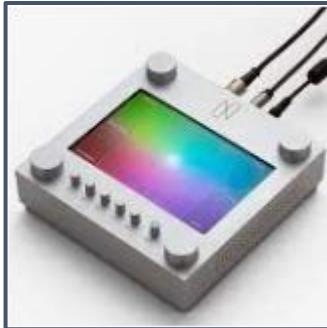
A2A Core Capabilities

- Customer service and support analysis
- Automating voice support
- Enhancing customer support qualities
- Creating media content and ads for energy sector customers.

Audio-to-Audio (A2A) Tools

ORGANIZER

ISGF
India Smart Grid Forum



NSynth Super

AudioLM



SpeechBrain



Enterprise Chatbot : A Case Study

ORGANIZER

- Enterprise wide single chatbot to handle different types of requests
 - Large Language Models(like GPT-3) tuned to use the knowledge articles and provide a precise answer to the user's query.
 - Enterprise Chatbot can handle any human interaction they have not been trained on, by using the powerful capabilities of LLMs, which come pre-trained on gigabytes of information from different useful sources.
 - The Chatbot can also be used to extract data from different applications for specific conditions by using APIs or the Natural Language to SQL convertor capability of the LLM, without having to hardcode any queries for the relevant application(s).
 - The LLM powered Chatbot can handle advanced requests like summarization and questions involving logical inference/comparison with respect to content/entities in multiple paragraphs within a document or spread across multiple documents.
- Includes the following Generative AI capabilities:
 - Text to Text
 - Text to Image
 - Speech to Text

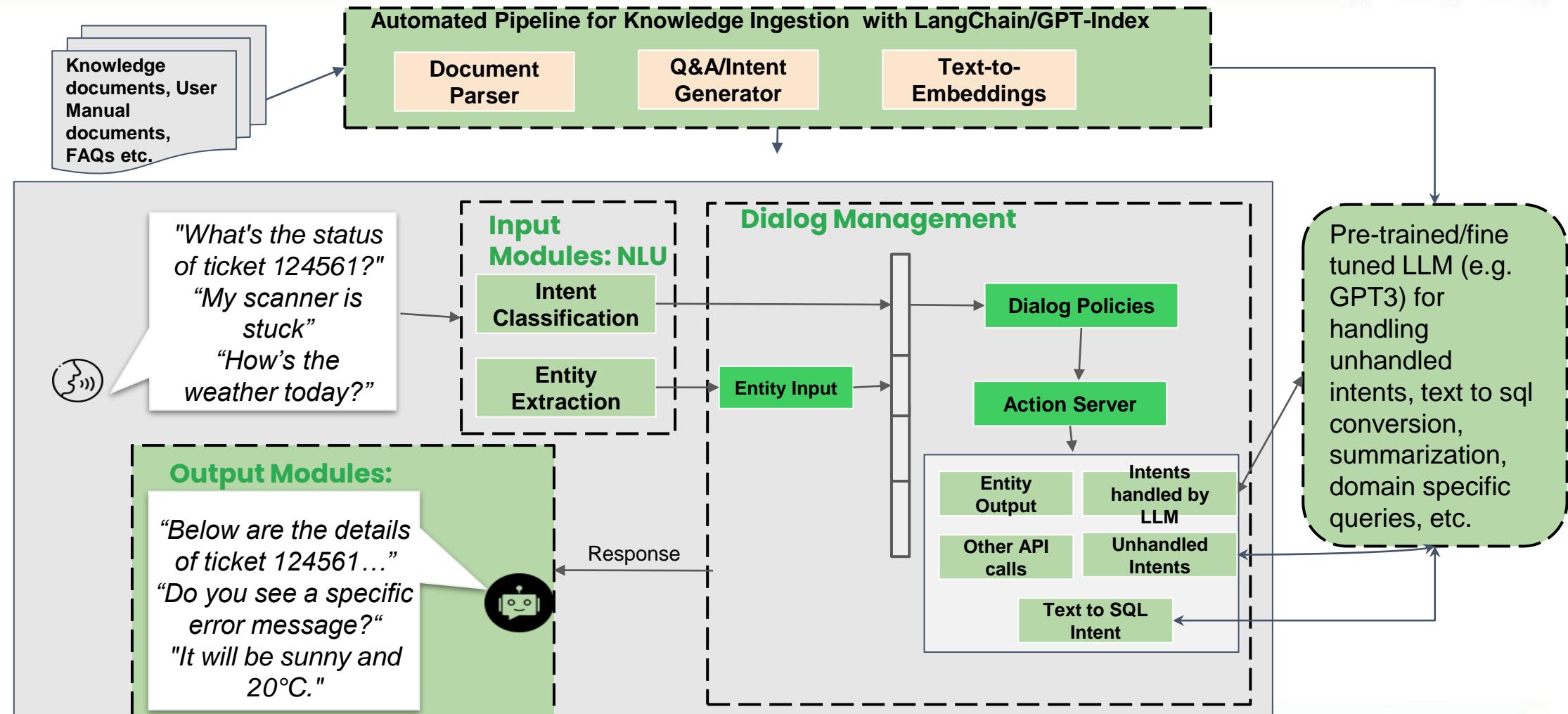
Enterprise Chatbot: Integrations

ORGANIZER

ISGF
India Smart Grid Forum

- Profile Management System
- Redmine
- Apache Superset
- Holiday Calendar
- Company Policy documents-
- General questions
- Dall-E-2

Enterprise Chatbot : Architecture



Thank You

For discussions/suggestions/queries email: isuw@isuw.in

www.isuw.in

[Links/References \(If any\)](#)