Edict.ai

Created with love by: Closed AI Chain

Table of contents

01 Introduction
Aim
Objectives
Problem statement

02 App's Journey
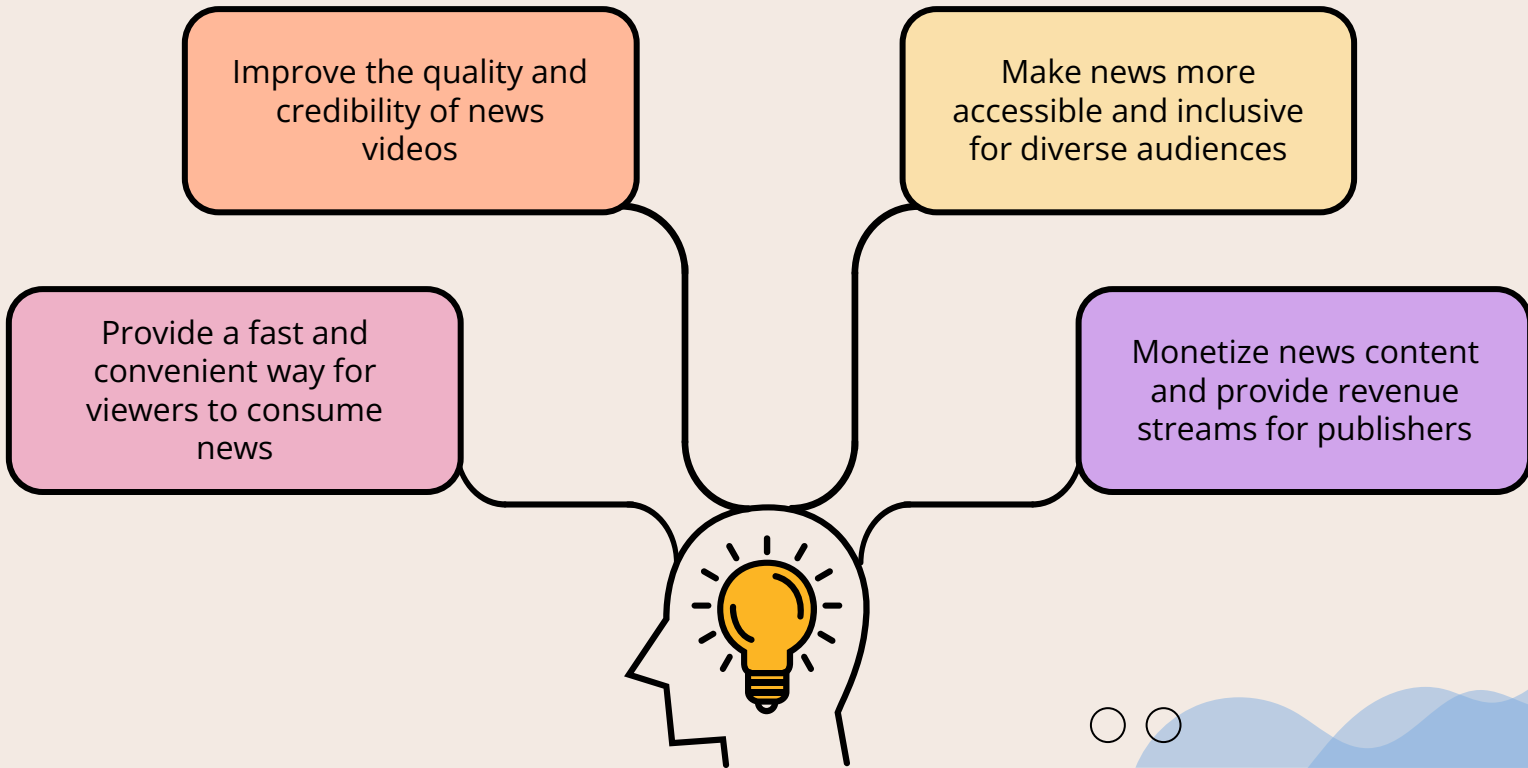System architecture
Phases

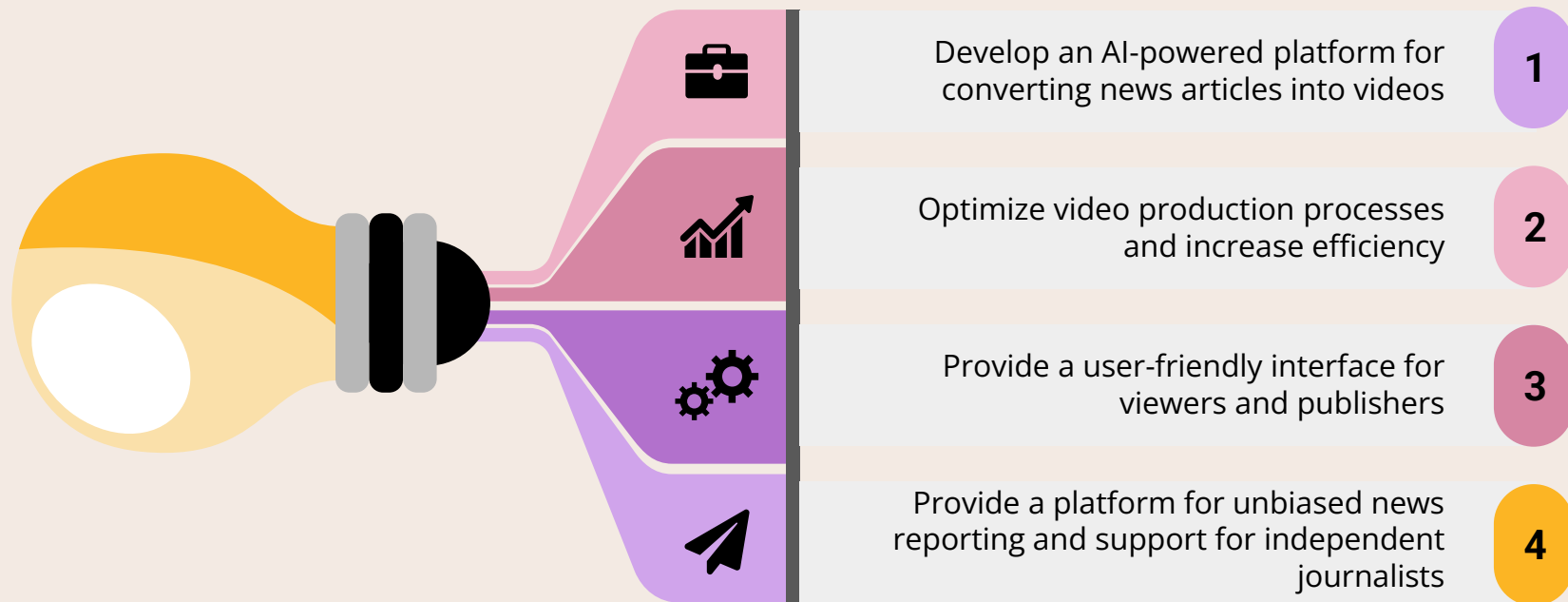**03 Business Model**
Usp
Market size
Revenue stream

04 Results
Conclusion
Future Scope

Aim

Improve the quality and credibility of news videos

Make news more accessible and inclusive for diverse audiences

Provide a fast and convenient way for viewers to consume news

Monetize news content and provide revenue streams for publishers

Objectives

1. Develop an AI-powered platform for converting news articles into videos

2. Optimize video production processes and increase efficiency

3. Provide a user-friendly interface for viewers and publishers

4. Provide a platform for unbiased news reporting and support for independent journalists
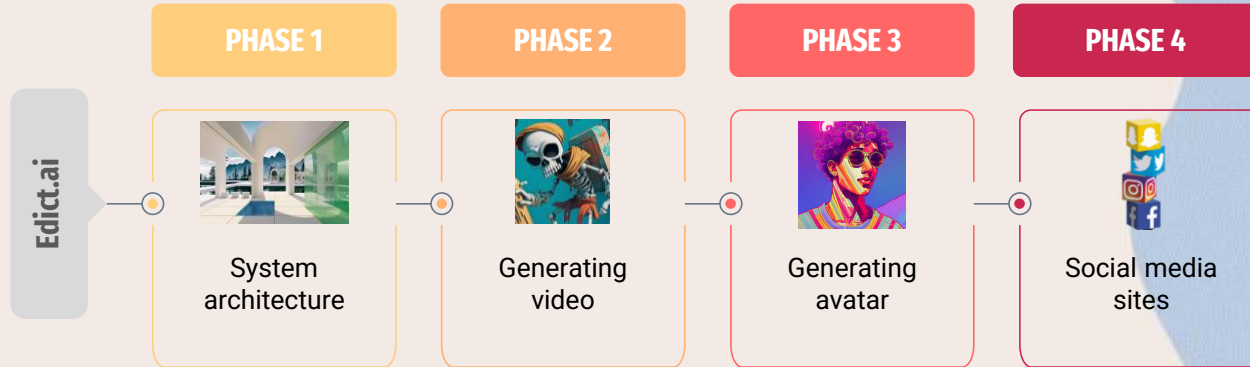
Problem Statement

Optimize video production workflows through automation and AI

Develop state-of-the-art natural language processing algorithms for script and voiceover creation
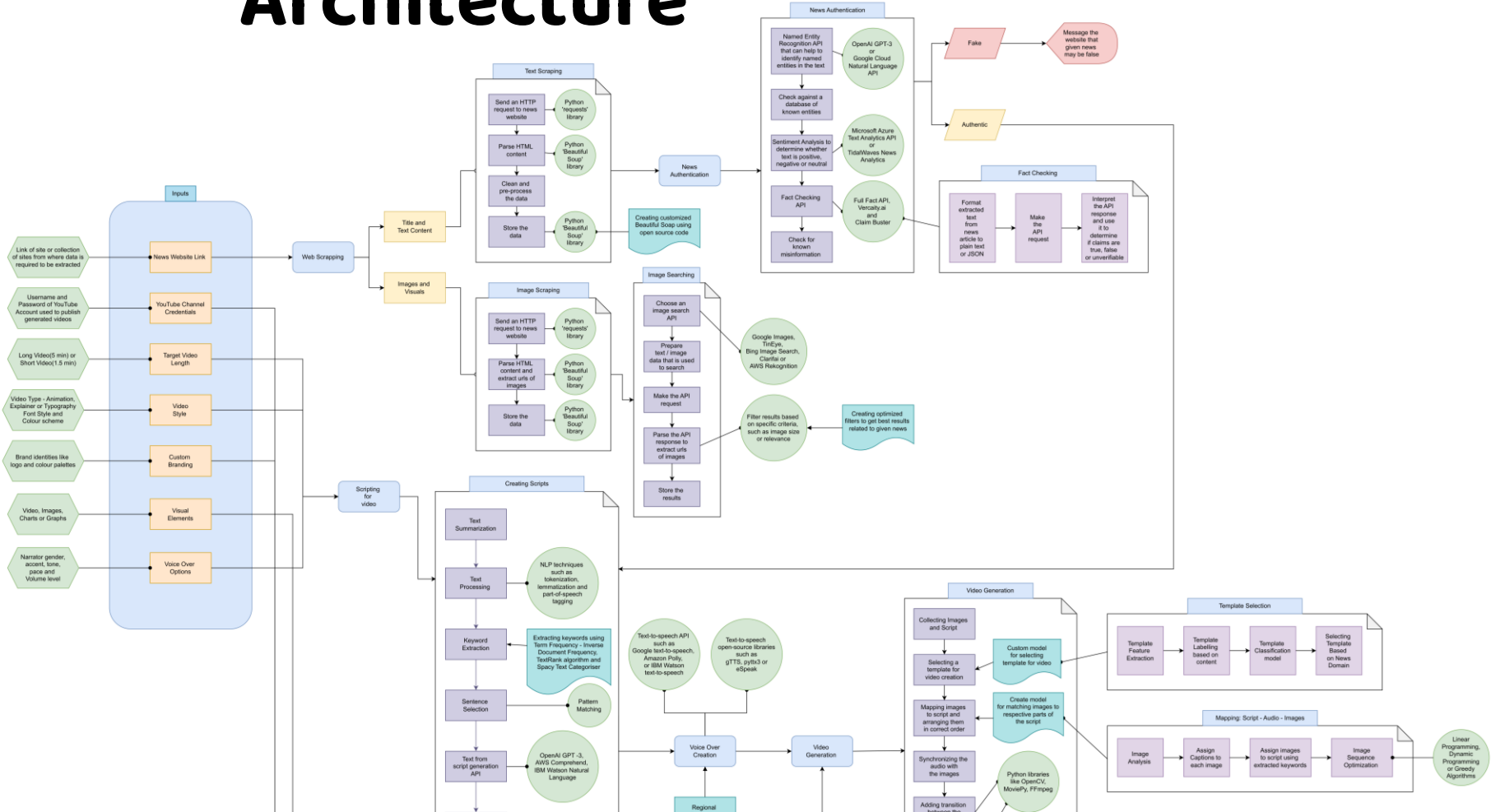
Build public trust in news reporting by providing transparent and unbiased news content
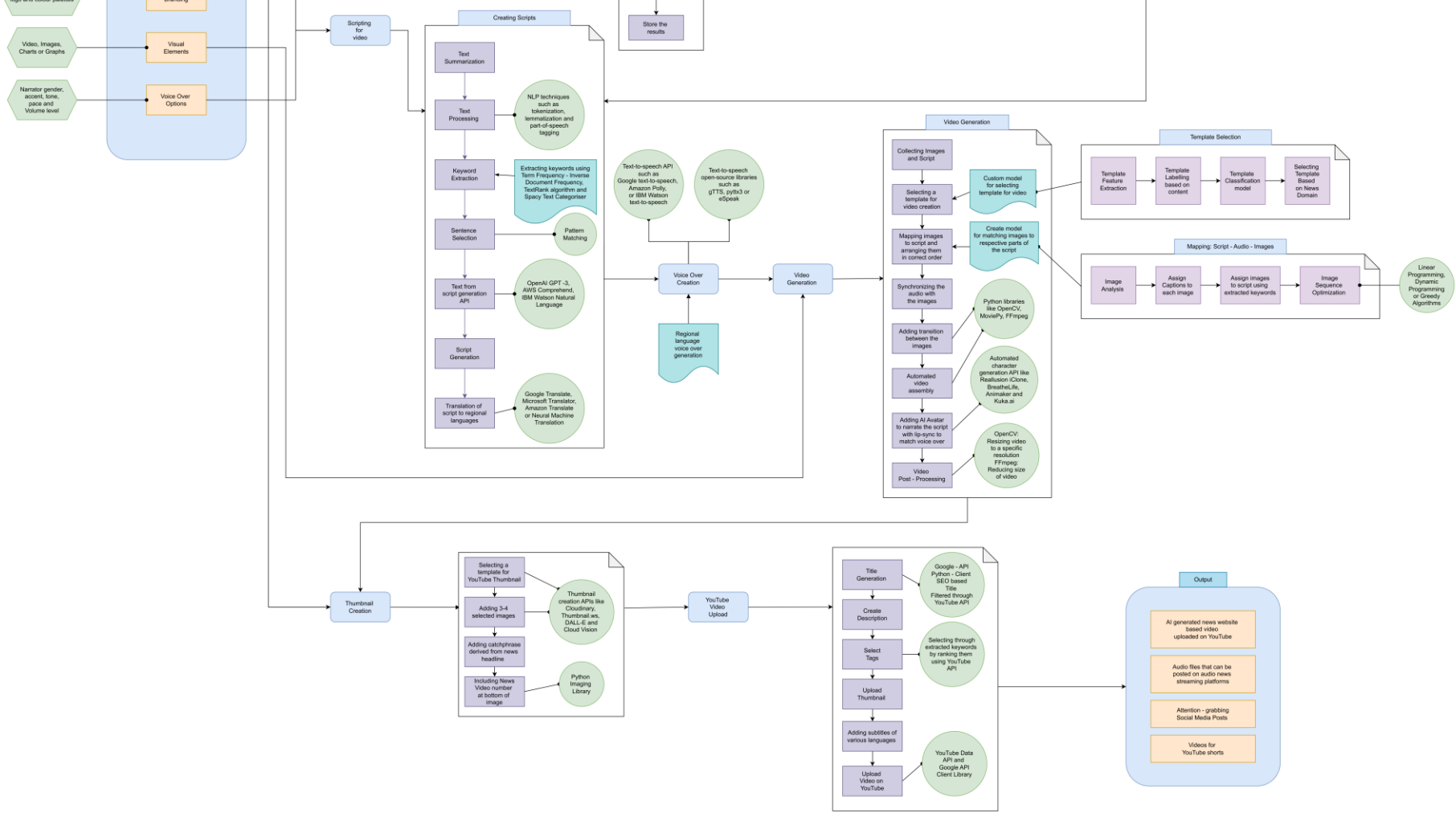
Establish partnerships with news publishers and advertisers to generate revenue
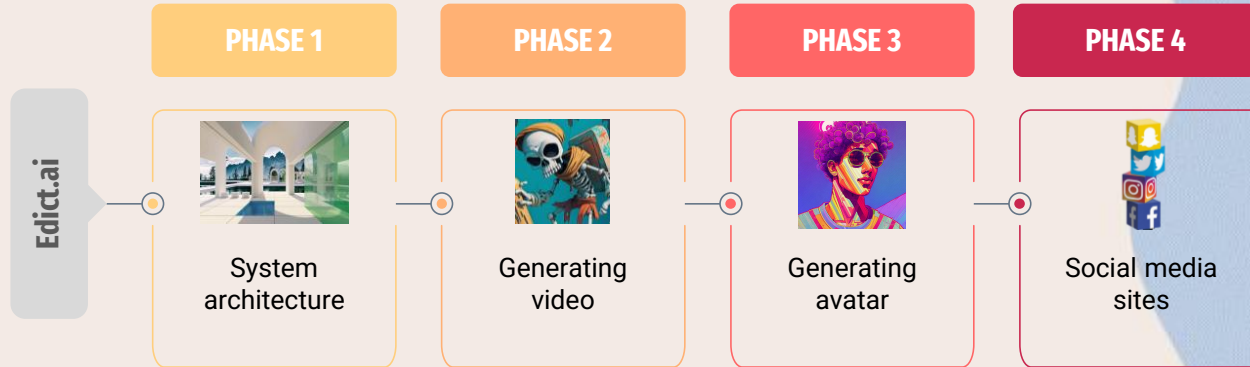
Evolution of edict.ai

| PHASE 1 | PHASE 2 | PHASE 3 | PHASE 4 |

Edict.ai

System architecture

Generating video

Generating avatar

Social media sites

Architecture

Video, Images, Charts or Graphs

Narrator gender, accent, tone, pace and Volume level

...Branding

Visual Elements

Voice Over Options

Scripting for video

Store the results

Creating Scripts

- Text Summarization
- Text Processing → NLP techniques such as tokenization, lemmatization and part-of-speech tagging
- Keyword Extraction → Extracting keywords using Term Frequency - Inverse Document Frequency, TextRank algorithm and Spacy Text Categoriser
- Sentence Selection → Pattern Matching
- Text from script generation API → OpenAI GPT -3, AWS Comprehend, IBM Watson Natural Language
- Script Generation
- Translation of script to regional languages → Google Translate, Microsoft Translator, Amazon Translate or Neural Machine Translation

Text-to-speech API such as Google text-to-speech, Amazon Polly, or IBM Watson text-to-speech

Text-to-speech open-source libraries such as gTTS, pytts3 or eSpeak

Voice Over Creation

Regional language voice over generation

Video Generation

Video Generation

- Collecting Images and Script
- Selecting a template for video creation → Custom model for selecting template for video
- Mapping images to script and arranging them in correct order → Create model for matching images to respective parts of the script
- Synchronizing the audio with the images
- Adding transition between the images → Python libraries like OpenCV, MoviePy, FFmpeg
- Automated video assembly → Automated character generation API like Reallusion iClone, BreatheLife, Animaker and Kuka.ai
- Adding AI Avatar to narrate the script with lip-sync to match voice over → OpenCV: Resizing video to a specific resolution FFmpeg; Reducing size of video
- Video Post - Processing

Template Selection

- Template Feature Extraction
- Template Labelling based on content
- Template Classification model
- Selecting Template Based on News Domain

Mapping: Script - Audio - Images

- Image Analysis
- Assign Captions to each image
- Assign images to script using extracted keywords
- Image Sequence Optimization → Linear Programming, Dynamic Programming or Greedy Algorithms

Thumbnail Creation

Thumbnail Creation group

- Selecting a template for YouTube Thumbnail
- Adding 3-4 selected images → Thumbnail creation APIs like Cloudinary, Thumbnail.ws, DALL-E and Cloud Vision
- Adding catchphrase derived from news headline
- Including News Video number at bottom of image → Python Imaging Library

YouTube Video Upload

- Title Generation → Google - API Python - Client SEO based Title Filtered through YouTube API
- Create Description
- Select Tags → Selecting through extracted keywords by ranking them using YouTube API
- Upload Thumbnail
- Adding subtitles of various languages
- Upload Video on YouTube → YouTube Data API and Google API Client Library

Output

- AI generated news website based video uploaded on YouTube
- Audio files that can be posted on audio news streaming platforms
- Attention - grabbing Social Media Posts
- Videos for YouTube shorts
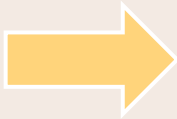
Scraping

Input	Process	Output
	• Accessing Title, Textual content and image urls through link	
	• Url Types - single article or multi-article	
URL		Title Content Image URLs
	Models and Libraries	
	Selenium library Beautiful Soup	

Authentication

Input	Process	Output
Scraped Content	• Checking if scrapped news is Fact checked or not • Checking if fake news is containing violence, hate speech type symptoms or not Models and Libraires: Decision Tree Classifier RapidAPI, Bind chat API	Authenticity (Real/Fake)

Scripting

Input	Process	Output
Scraped Content	1. Summarization of scrapped content 2. Adding meaningful and pause based Punctuation to create script 3. Chunking 4. Extraction of keywords from each chunk Models and Libraries: PeagasusCNN Daily Mail, Xsum-large,newsroom Keybert, BERT-base-ner	• Meaningful Chunks of scraped content • Keywords for every chunk
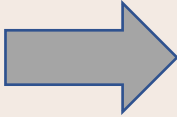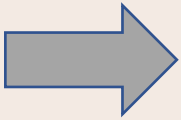
Image Search

Input

Keywords for every chunk

Process

Searching for images over keywords extracted from each chunk and saving them in particular folder.

Model and Libraries:
Google image Search API, Bing Image search api

Output

Image for keyword of that chunk
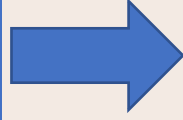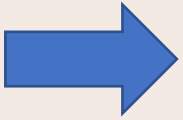
Audio
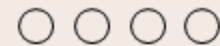
Input	Process	Output
Chunks	• TTS is an Google Text-to-speech API which convert text to speech. • SpeechBrain/TTS-Tacotron2-lgspeech • Microsoft /SpeechT5_TTS • Facebook/FastSpeech2-en-ljspeech Models: TTS , Tacotron	Audio for Chunk
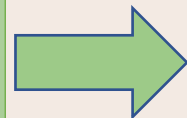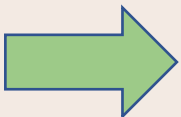
Merging
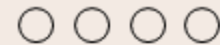
Input

chunks, images, intro, outro, template

Process

1. Converting image to video of for a particular audio chunk duration
2. Adding audio
3. Adding overlaying chunk_video template
4. Adding chunks as text subtitles
5. Exporting chunk
6. Merging all chunks to create final_video by adding intro and outro

Model and Libraries:-
Moviepy

Output

Final Video Generated
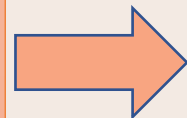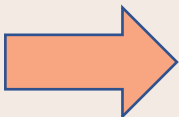
Thumbnail

Input

1
Background
Main
Image,2
sub-images
Important
Headings

Process

1. Using Pillow library setting background

2. Adding two sub-images at particular positions

3. Adding headings

Model and Libraries:

Pillow,random2

Output

1280x720
Size Youtube
Thumbnail
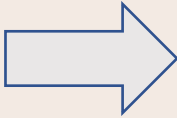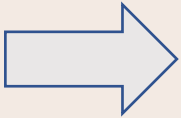
YouTube Upload with SEO Optimization
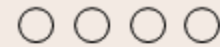
Input

Video, Youtube API, Secret_client credential, scrapped title, thumbnail

Process

1. Authentication of Youtube Channel with google-auth-python-api library, google-ouath2
2. Getting SEO based tags and description
3. Uploading video with title, thumbnail, description, tags

Model and Libraries:

Youtube V3 Data API, google-oauth2, google-auth-python-api

Output

Video Uploading in Youtube

Evolution of edict.ai

| PHASE 1 | PHASE 2 | PHASE 3 | PHASE 4 |

Edict.ai

System architecture

Generating video

Generating avatar

Social media sites
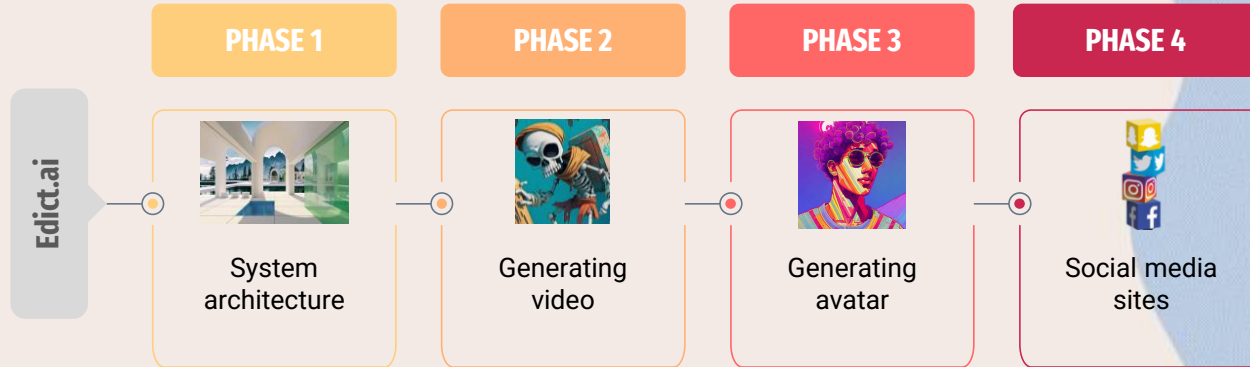
Phase 4: Future of edict.ai

NLP + Ocr + image processing + pdf analysis

Blockchain consensus mechanism, Backtracking sources, 4 model architecture

Multilanguage support, Choice of voice, Ssml

Reverse Image Search, Cosine Similarity

Merging with AI Avatar, Custom intro & outro, Template selection

Multiplatform uploading support, shorts

Literature Survey

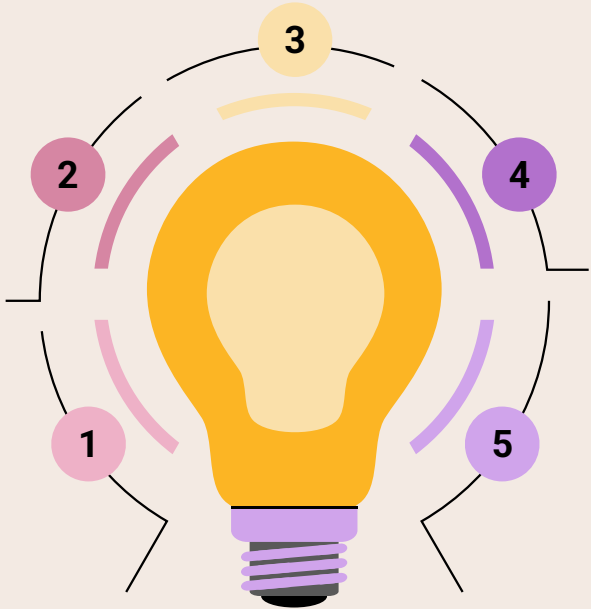| Name	Challenges	Proposed Solution	Research Gap
"An Approach of Web Scraping on News Website based on Regular Expression," [2018]	1. Non-Uniform Website Structure 2. IP blocking 3. Security pass 4. Data filtration 5. Updation of scrapping pattern as per web updates	1. Filter patter method 2. Pandas, Polars 3. VPN can be used to bypass security and extract more data	Web scraping tools struggle to automatically detect changes in website structure and update scraping code patterns, which can result in outdated or incomplete data
A deep convolutional neural network for fake news detection, Cognitive Systems Research, [2020]	1. Contextual Understanding 2. Multi-modal detection – fake news can be spread in multiple forms media, text, audio 3. Multi-lingual and cross cultural	1. Multimodal frameworks detects fake news in different forms 2. NLP, Graph, deep learning	1. Developing AI can generate fake videos which can mislead authentication models in identification of fake news. 2. Detection for multilingual news is not much explored

NAME	Challenges	Proposed Solution	Research Gap
Building Multilingual End-to-End Speech Synthesizers for Indian Languages [2019]	1. Human tone 2. Multilingual 3. Multi-dialect	1. Multilingual speech synthesis model is discussed	Multi-dialect speech synthesis
From Show to Tell: A Survey on Deep Learning-Based Image Captioning [2022]	1. Meaningful mapping 2. Duration of image should be matched with script/audio	Image captioning can help to do meaningful mapping	Complexity of model is high to match the image duration accurately with voiceover
Automatic generation of visual-textual web video thumbnail [2017]	1. Multiple domains 2. Creativity, Quality and testing 3. Thumbnails should not contain text or media which is misleading or inappropriate 4. Copyright issues	1. Algorithms can help model to understand the context 2. Layout Formation	1. Single multi-domain thumbnail creation 2. Model to authenticate that thumbnail should not contain misleading or clickbait objects 3. Limited flexibility to generate different viewer engaging layouts

Conclusion

High potential in tems of Technical feasibilty

3

Democratize the news industry and empower independent journalists

2

Significant business potential

4

Revolutionizing the way news content is consumed

1
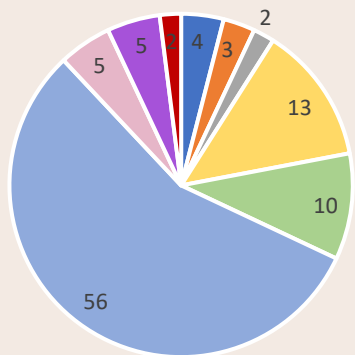
Project has the potential to create a more inclusive and diverse landscape of news content for audiences around the world

5

Results

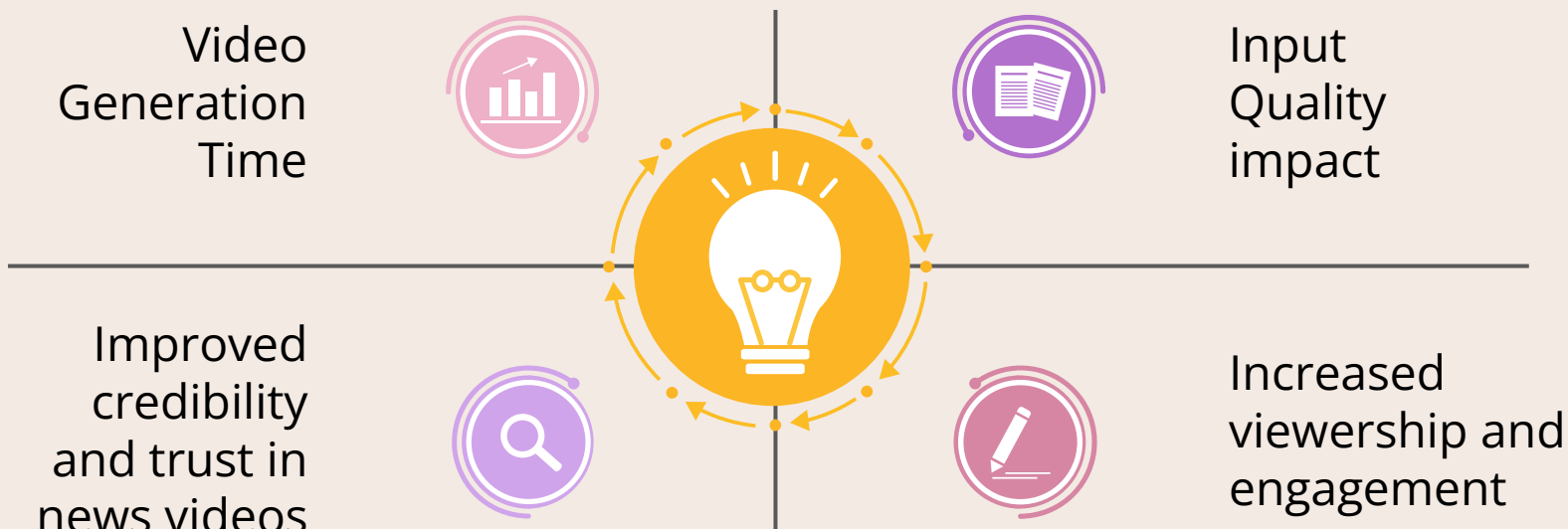## Video generation time distribution

Web scraping	4	16
News Authentication	3	15
Scripting	2	5
Image searching	13	44
Avatar and audio generation	10	35
Chunk generation	56	205
Overlaying avatar on chunk	5	20
Merging video	5	22
Thumbnail and YouTube video uploading	2	10

Legend:
- Web scraping
- News Authentication
- Scripting
- Image searching
- Avatar and audio generation
- Chunk generation
- Overlaying avatar on chunk
- Merging video
- Thumbnail and YouTube video uploading

Results

Video Generation Time

Input Quality impact

Improved credibility and trust in news videos

Increased viewership and engagement

Market Size

Digital Newspapers & Magazines Market Report Scope

Report Attribute	Details
Market size value in 2022	USD 35.60 billion
Revenue forecast in 2028	USD 45.75 billion
Growth rate	CAGR of 4.2% from 2022 to 2028
Base year for estimation	2021

The global digital video content market size reached US$ 171.5 Billion in 2022.

Looking forward, the publisher expects the market to reach US$ 355.2 Billion by 2028, exhibiting a CAGR of 12.9% during 2022-2028

USP

Automated News-to-Video Conversion

Seamless Social Media Integration

Customization and Personalization

Cost and Time Efficiency

SEO Optimization and Traffic Generation

Revenue Streams

Premium Subscription:

    :Will be providing more features such as

    add their on photo as a Avataar (talking head) in the videos.

    Customization: Adding Intros and outros

    Enhanced Visual Effects

    Can use our APIs

    Ad free Videos

    SEO Optimization
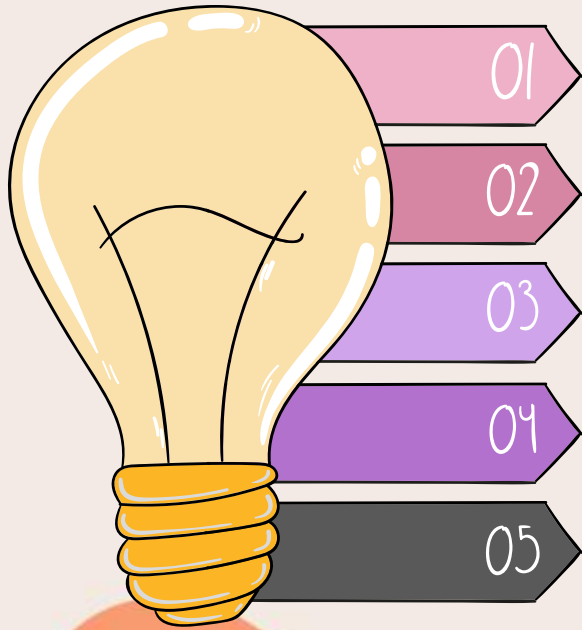
Advertising And Monitization:

    We will be displaying adds on the website.

    We can add ads in the News Videos also.

Providing Paid APIs

    We can provide Paid APIs for organizations who want t integrate our service on their platform.

References

01 A. Maududie, W. E. Y. Retnani and M. A. Rohim, "An Approach of Web Scraping on News Website based on Regular Expression," 2018 2nd East Indonesia Conference on Computer and Information Technology (EIConCIT), Makassar, Indonesia, 2018, pp. 203-207, doi: 10.1109/EIConCIT.2018.8878550.

02 Rohit Kumar Kaliyar, Anurag Goswami, Pratik Narang, Soumendu Sinha, FNDNet – A deep convolutional neural network for fake news detection, Cognitive Systems Research, Volume 61, 2020, Pages 32-44, ISSN 1389-0417, https://doi.org/10.1016/j.cogsys.2019.12.005

03 Prakash, Anusha & Thomas, Anju & Umesh, Santhosh & Murthy, Hema. (2019). Building Multilingual End-to-End Speech Synthesisers for Indian Languages. 194-199. 10.21437/SSW.2019-35

04 Stefanini, Matteo & Cornia, Marcella & Baraldi, Lorenzo & Cascianelli, Silvia & Fiameni, Giuseppe & Cucchiara, Rita. (2022). From Show to Tell: A Survey on Deep Learning-Based Image Captioning. IEEE Transactions on Pattern Analysis and Machine Intelligence. PP. 1-1. 10.1109/TPAMI.2022.3148210

05 Zhao, Baoquan & Lin, Shujin & Qi, Xin & Zhang, Zhiquan & Luo, Xiaonan & Wang, Ruomei. (2017). Automatic generation of visual-textual web video thumbnail. 1-2. 10.1145/3145690.3145712