UiPath Reboot Work.™ | Process Definition Document

YouTube Comment Filter

Our YouTube Comment Filter project is designed to streamline the comment analysis process on YouTube videos. By selecting a video, the system performs scraping on the first chosen number of comments. Leveraging an Excel mapping system and artificial intelligence, the comments are categorized into "Positive," "Negative," and "Neutral". In the end, the evaluation results are efficiently compiled and sent via email, providing a comprehensive overview of the community's feedback on the video. This automation enhances efficiency and allows content creators to better understand and engage with their audience.

Contents

1. Introduction

1.1 Purpose

The document defines the selected business process for automation using UiPath Robotic Process Automation (RPA) technology. It outlines the sequence of steps performed as part of the process, the conditions and rules of the process before automation, and how they are expected to operate after its partial or complete automation. This document provides the necessary details for RPA developers to implement the robotic automation of the chosen business process.

1.2 Objectives

The business objectives and benefits expected by the Business Process Owner after the implementation of the YouTube Comment Filter project are:

- Enhanced Efficiency in Comment Analysis: By automating the process of scraping and categorizing the first chosen number of comments from YouTube videos, the system significantly speeds up the comment analysis process. This allows for a quick and efficient overview of public sentiment, reducing the time and effort required compared to manual analysis.
- Streamlined Categorization through AI: Utilizing an Excel mapping system and artificial intelligence, comments are automatically categorized into "Positive," "Negative," and "Neutral." This sophisticated categorization aids in better understanding the community's feedback, ensuring that content creators can easily identify general sentiments and respond accordingly.
- Speed of Comment Processing: The automated system can scrape and categorize the first chosen number of comments from a YouTube video almost instantaneously. In contrast, manual analysis of the same number of comments could take hours, depending on the complexity and length of the comments.

1.3 Key Contacts

Role	**Name**	**Contact Details** (email, phone number)	**Notes**
Developer	**Mogage Nicolae**	**nicolae.mogage@stud.ubbcluj.ro 0747081285**	
Developer	**Marian Daria-Georgiana**	**daria.marian@stud.ubbcluj.ro 0748165858**	
Developer	**Martin Fabian-Ionuț**	**fabian.martin@stud.ubbcluj.ro 0745138776**	
Developer	**Molnar Noemi-Eveline**	**noemi.eveline.molnar@stud.ubbcluj.ro 0727296868**	
Developer	**Matei Otniel-Daniel**	**otniel.matei@stud.ubbcluj.ro 0763307843**	

1.4 Minimum Pre-requisites for the Automation

a) Filled in Process Definition Document
b) Test Data to support development
c) Valid Email for Communication: Users must provide a valid email address. This email will be used to send the compiled analysis of the YouTube comments, enabling a direct line of communication and feedback.

d) Valid YouTube Link: Users need to provide a valid YouTube link for the video they wish to analyze. The system uses this link to scrape the first chosen number of comments for analysis.
e) Robust Internet Connection: A stable and robust internet connection is crucial for scraping comments in real-time and ensuring uninterrupted processing.

2. AS IS Process Description

2.1 Process Overview

Item	Description/Answer
Process Full Name	***YouTube Comment Filter***
Process Area	Comment Classification
Department	G02
Short Description (operation, activity, outcome)	The user needed to open a search engine, then YouTube, choose their desired video to analyze and then copy each comment individually and put it in an excel worksheet. From there, they would have put each comment on one of the three categories (bad, neutral, good) according to their own reasoning.
Role(s) required in applications to perform the process	Solution Architect, Business Analyst, Developer
Process schedule and frequency	3 videos per day
Number of times the process is ran by selected frequency	1
Process execution time	***>1 hour***
Process Restrictions	*n/a*
Peak Period (s)	***n/a***
Peak Volume Approximate increase	***n/a***
Number of persons performing the process	***1***
Expected Volume increase during next periods	*n/a*
Percentage Un-handled exceptions	***unknown***
Input data description	*YouTube video link*
Output Data description	*Excel worksheet*

2.2 Applications Used

Application Name	Version	Application Language	Thin/Think Client	Environment/ Access method	Comments
Excel	latest version included in the Microsoft 365 package	VBA	Thick Client	On-premises	Automated data analysis
Chrome	120.0.6099.199	English	Thin Client	Remote desktop	Default web browser
OpenAI	GPT-3.5	Python	-	API	Natural language processing

2.3 Process Statistics

High Level statistics

Processes	Windows	Actions	Mouse clicks	Keys pressed	Text entries	Hotkeys used	Time
1	2	100+*	100+*	100+*	100+*	Ctrl+c, ctrl+v	>60min

Detailed statistics

Window name	Mouse clicks	Text entries	Key pressed
Chrome	3	Video name - 1	Letters + enter
Excel	100+*	100+*	100+*

** This varies depending on the number of comments that the user chooses to analyze. If he chooses 100, and he enters 3 links, it could be 300+, but not necessarily (maybe not all of the videos have 100 comments). Or if he chooses to get all the comments analyzed, there is no certainty of the exact number.*

2.4 Detailed AS IS Process Actions

#Action	Input	Description	Details (Screen/Video Recording Index	Exception Handling	Possible Actions
Open a Search Engine	Web browser	Accessing a search engine	User opens their web browser and navigates to a search engine like Google, Bing, etc.	If the search engine is down	Try a different search engine or check internet connectivity
Navigate to YouTube	Search engine	Locating YouTube	User types 'YouTube' into the search engine or enters the YouTube URL directly.	If YouTube is inaccessible	Check internet connection, try accessing later
Search for Desired Video	YouTube platform	Finding a specific video	Using YouTube's search bar, user	If video is not found	Refine search keywords,

#Action	Input	Description	Details (Screen/Video Recording Index	Exception Handling	Possible Actions
			searches for the video they want to analyze.		ensure correct spelling
Open Selected Video	Search results	Opening the video page	User clicks on the desired video from the search results to open it.	If video won't play	Refresh the page, check internet connection
Manual Comment Collection	YouTube video page	Copying comments individually	User scrolls through comments, manually copying each one.	If comments are disabled	Choose another video
Open Excel Worksheet	Computer, Excel software	Preparing to organize comments	User opens an Excel worksheet on their computer.	If Excel is not available	Use alternative spreadsheet software
Paste Comments into Excel	Copied comments	Transferring comments to Excel	User pastes each copied comment into individual cells in Excel.	Errors in copying/pasting	Re-copy the comment, check for formatting issues
Manual Categorization	Excel worksheet with comments	Sorting comments into categories	User reads and categorizes each comment as bad, neutral, or good based on their judgment.	Uncertainty in categorization	Use a set of guidelines for consistent categorization
Compile Results	Categorized comments	Reviewing and summarizing findings	User reviews categorized comments to compile analysis results.	Inconsistencies in data	Recheck categorizations, ensure accurate counting
Report Generation	Analysis results	Creating a summary of the categorization	User may create a summary report or visual representation of the categorization results.	Difficulty in report creation	Use report templates, seek assistance if needed

2.5 Input Data Description

#Action	Sample	Input Type	Location	Are inputs Natively Digital*?	Are the inputs Structured*?
Open a Search Engine	-	Web browser interface	User's computer	Yes	Yes
Navigate to YouTube	Web browser	URL or search query	User's computer	Yes	Yes

#Action	Sample	Input Type	Location	Are inputs Natively Digital*?	Are the inputs Structured*?
Search for Desired Video	YouTube platform	Text query	YouTube search bar	Yes	No
Open Selected Video	Search results	Clickable video link	YouTube page	Yes	Yes
Manual Comment Collection	YouTube video page	Video comments	Below YouTube video	Yes	No
Open Excel Worksheet	Computer, Excel software	Excel software interface	User's computer	Yes	Yes
Paste Comments into Excel	Copied comments	Text (comments)	Excel sheet	Yes	No
Manual Categorization	Excel worksheet with comments	Text (comments) for review	Excel sheet	Yes	No
Compile Results	Categorized comments	Categorized data	Excel sheet	Yes	Yes
Report Generation	Analysis results	Summarized data, charts, etc.	Excel or other tools	Yes	Yes

** Native Digital: This is data that was originally created digitally e.g. excel, database or application reports etc. The non-native digital inputs are usually scanned images.*

** Structured Data: has a predictable format and exists in fixed fields (e.g. an excel cell or a field in a form) and is easily detectable via search algorithms.*

3 TO BE Process Description

3.1 Detailed TO BE Process Map

Action Number in the Process	Process Action	Proposed for Automation?	Remains Manual?
1	Access YouTube and Locate Desired Video	Yes	No
2	Scrape First *maxResults* Comments from the Video, where *maxResults* = the maximum number to analyze (chosen by the user; the user can also choose to analyze all of the comments)	Yes	No
3	Categorize Comments into Sentiments	Yes	No
4	Compile Categorized Data into a Report	Yes	No
5	Analyze Data for Key Themes and Insights	Yes	No
6	Generate Comment Analysis Report	Yes	No
7	Send Report via Email to User	Yes	No
8	Monitor System Performance and Log Data	Yes	No
9	Review and Address System Errors/Exceptions	No	Yes
10	Collect User Feedback and Suggestions	No	Yes

3.2 Parallel Initiatives

Initiative Name	Process Action(s) where it is identified	Impact on current Automation Request	Expected Completion Date	Contact Person
Upgrade OpenAI API Version	Any action involving OpenAI interactions	May require adjustments in OpenAI-related workflows. Ensure compatibility with the new API version.	Q2 2024	Team
Implement Multilingual Support	Text processing and language-specific actions	Enhance the robot's capability to process comments in multiple languages. May impact language-related modules.	Q3 2024	Team
Integration with New Analytics Tool	Data extraction and analysis steps	Improve data analytics capabilities. Automation may need to adapt to new data formats or sources.	Q4 2024	Team
User Interface Redesign	Any action involving UI interactions	Interface changes may affect UI automation. Adjustments may be needed for screen element recognition.	Q1 2025	Team

3.3 In Scope for RPA

- **YouTube Video Selection:**

Identify and select a YouTube video for comment analysis.

- **Web Scraping for Comments:**

Scrape the first *maxResults* Comments from the Video, where *maxResults* = the maximum number to analyze (chosen by the user; the user can also choose to analyze all of the comments)

- **Comment Categorization:**

Utilize Excel mapping and AI to categorize comments into "Positive," "Negative," and "Neutral."

- **Results Compilation:**

Compile the results of comment categorization.

- **Email Notification:**

Send an email with the evaluation results.

- **Exception Handling:**

Implement error-handling mechanisms for potential exceptions during the process.

- **User Interface Interactions:**

Perform necessary UI interactions, if applicable, for video selection and other user-related actions.

- **Data Validation:**

Validate data inputs and outputs to ensure accuracy and reliability.

3.4 Out of Scope for RPA

Activity/Action*	Reason for out of scope	Impact on the TO BE	Possible measures to be taken into consideration for future automation
Handwritten Data Entry	Input is handwritten	Manual data entry required	Collect input in a digital format or use OCR for handwritten text recognition.
Highly Dynamic Web Pages	Frequent UI changes	Automation script may break	Implement a periodic review and update process for UI elements in future automation.

3.5 Exception Handling

3.5.1 Known Business Exceptions

Exception Name	Action	Parameters	Actions to be taken
BusinessRuleException ("Too many emails!")	*numberOfValidEmails.Count > 1*	*numberOfValidEmails*	*Show a message box with text "You added more than 1 email. Bye :(" and rethrow the exception. The application stops.*
BusinessRuleException("Too many videos!")	*linkCount > 3*	*linkCount*	*Show a message box with text "You added more than 3 videos. Bye! :(" and rethrow the exception. The application stops*

3.6 Applications Errors & Exceptions Handling

3.6.1 Known Applications Errors and Exceptions

Error/Exception Name	Action	Parameters	Actions to be taken
OpenAI Rate Limit Exceeded	*Any OpenAI-related action*	*Error message: "429 Too Many Requests"*	*Recover and retry 3 times with appropriate back-off intervals.*

3.7 Reporting

Report Type	Update frequency	Details	Monitoring Tool to visualize the data
YouTube Comment Filter - Analysis Report	After each analysis	From: raportgenerationemail@gmail.com To: [Recipient's email] Subject: YouTube Comment Filter - Analysis Report Body: Dear [Recipient's email], We hope this email finds you well. Attached, you'll find the analysis report for the video/s: [Bullets for video(s) name(s)] Should you have any questions or require further information, please don't hesitate to reach out. Best regards, Dev Team Contact us: raportgenerationemail@gmail.com	UiPath, Excel, SMTP

4 Other

4.1 Additional sources of process documentation

Additional Process Documentation		
Video Recording of the process	Normal flow	https://www.youtube.com/watch?v=PnRUnP_1Nig
	Business exceptions flows	https://youtu.be/fxz8f5Agv58