Supplementary Material for SmartFlow: Robotic Process Automation using LLMs

Arushi Jain, Shubham Paliwal, Monika Sharma, Lovekesh Vig, Gautam Shroff
TCS Research
New Delhi, India
Email:{j.arushi@tcs.com}

ACM Reference Format:
Arushi Jain, Shubham Paliwal, Monika Sharma, Lovekesh Vig, Gautam Shroff. 2023. Supplementary Material for SmartFlow: Robotic Process Automation using LLMs. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (CIKM'23).* ACM, New York, NY, USA, 4 pages. https://doi.org/XXXXXXX.XXXXXXX

1 ADDITIONAL FIGURES

Due to page-limit of the paper, we provide figures which can easily convey the idea of the paper in this section. Figure 1 illustrates the variations in the layouts of different applications. The proposed SmartFlow RPA is designed to be adaptable to different layouts or applications without any manual customization.

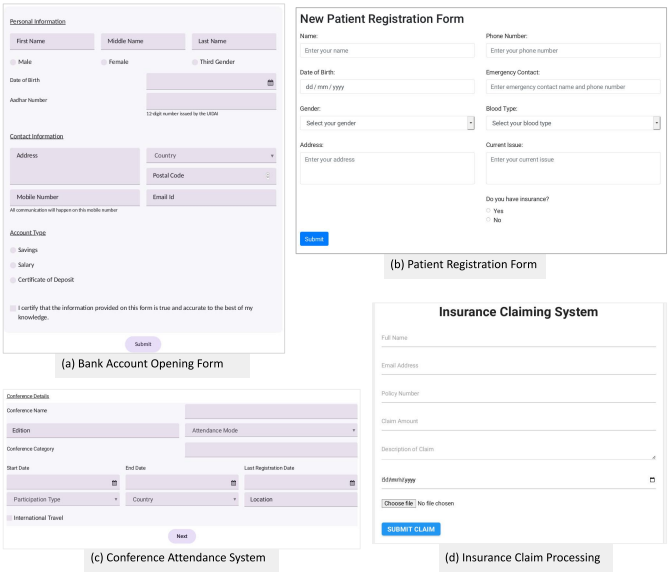

Figure 1: Figure showing examples of diverse application layouts.

Next, we provide an overview of the problem that we are trying to solve with SmartFlow in Figure 2. It is followed by an overview of our proposed approach in Figure 3. We also illustrate the process of layout mapping in Figure 5 and Figure 5.

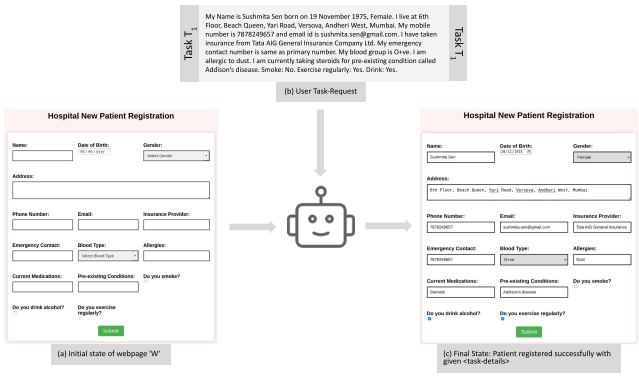

Figure 2: Given a user task-request to register a new patient on the hospital website W with given details as T_1, SmartFlow performs the task automatically and returns the status as "Patient registered successfully".

2 HANDLING COMPLEX COMPONENTS

Smartflow, an AI-based RPA system, has been designed to handle complex field types such as date pickers, dropdown menus, upload buttons, radio buttons, and checkboxes with seamless ease. For this purpose, individual algorithms based on vision techniques have been developed for each category of fields, as discussed below:

- *Datepickers*: They are characterized by their diverse and captivating designs, pose a unique challenge in the realm of user interaction. They are characterized by their diverse and captivating designs. In the majority of the layouts encountered, datepickers can be effortlessly populated by simply typing the relevant date. However, in scenarios where manual input is restricted, SmartFlow employs an astute strategy to ensure optimal performance. It intelligently scrolls through the corresponding year section by using PyAutoGUI and visually analyzing the screenshot of the calendar and hence, diligently seeks to match the desired year, adapting its direction based on whether the year lies in the past or future. It further refines the selection process by meticulously clicking through each subsequent month until the desired month aligns harmoniously. With the calendar now seamlessly set, Smartflow

CIKM'23, Oct 21–25, 2023, Birmingham, UK

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

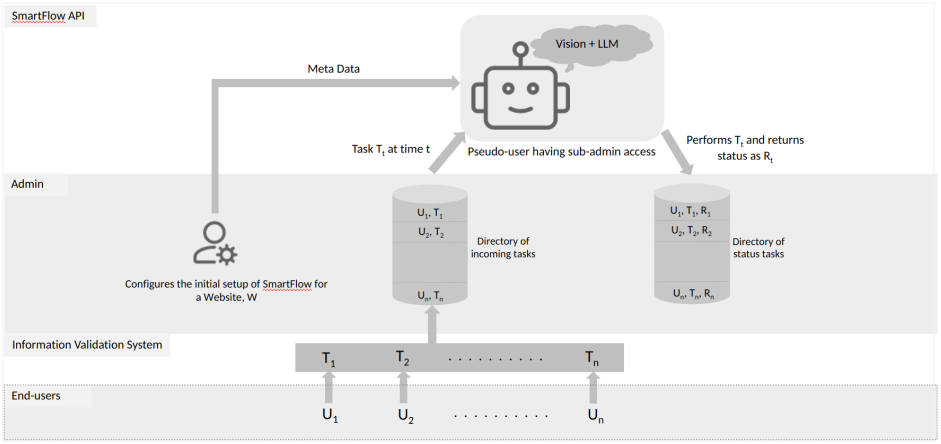

Figure 3: The proposed system overview of SmartFlow RPA: The process begins with end users (U_1, U_2,...,U_n) submitting corresponding task-requests (T_1, T_2,...,T_n) to fill out a form in a web application, which must include all mandatory details. These requests are then validated by an Information Validation System (IVS) and are added to a directory of incoming tasks for processing by the Smartflow system. As a one-time setup, the administrator specifies the URL of the website where end user requests are executed, along with the required HTML source code of the application pages and corresponding layout mapping of the input fields. The SmartFlow executes each task-request using machine vision and natural language understanding techniques, returning the status of the completed tasks to the directory of task status.

harnesses the power of vision and a simple technique of the calendar calculation to extrapolate the coordinate for each date, thus enabling flawless and accurate selection.

- *Dropdown*: SmartFlow, our proposed RPA system, effectively handles complex dropdown menus that have initially hidden options by leveraging visual processing techniques. When it encounters a <select> field, it clicks on the designated area of a dropdown menu, captures a screenshot to eliminate visual discrepancies, and extracts the dropdown options as digital representations using a text extraction module. In cases where the desired selection is not immediately visible, SmartFlow intelligently scrolls the dropdown panel until it locates the target selection, ensuring a smooth and seamless user experience.
- *Radio-buttons / Checkboxes*: To handle radio-buttons / checkboxes, Smartflow recognizes that it doesn't require clicking directly on the checkbox or circle. It utilizes the text associated with the option for selection. Smartflow interacts with ChatGPT to gather the available options and coordinates, and a dedicated module performs the action of clicking on the desired option from the list.

These algorithms for handling complex fields are invoked by the LLM during the generation of the navigation workflow. They are premised on a comprehensive understanding of the visual aspects of the form and integrate nuanced insights into the behavioral patterns of each field type. This ensures the creation of robust code blocks that can handle these fields effectively.

3 EXPERIMENTAL RESULTS

In this section, we provide detailed evaluation results of SmartFlow on RPA-dataset applications such as Patient Registration, Sales Lead Generation, Customer Complaint and Passport Registration applications in Tables 1, 2, 3, 4, respectively.

REFERENCES

Figure 4: Figure illustrating Layout Mapping using Demonstration by Admin approach. This approach utilizes visual understanding techniques to determine the layout mapping by providing an empty state, a filled state, and a JSON file containing field names, placeholders/edit fields, and their corresponding field values.

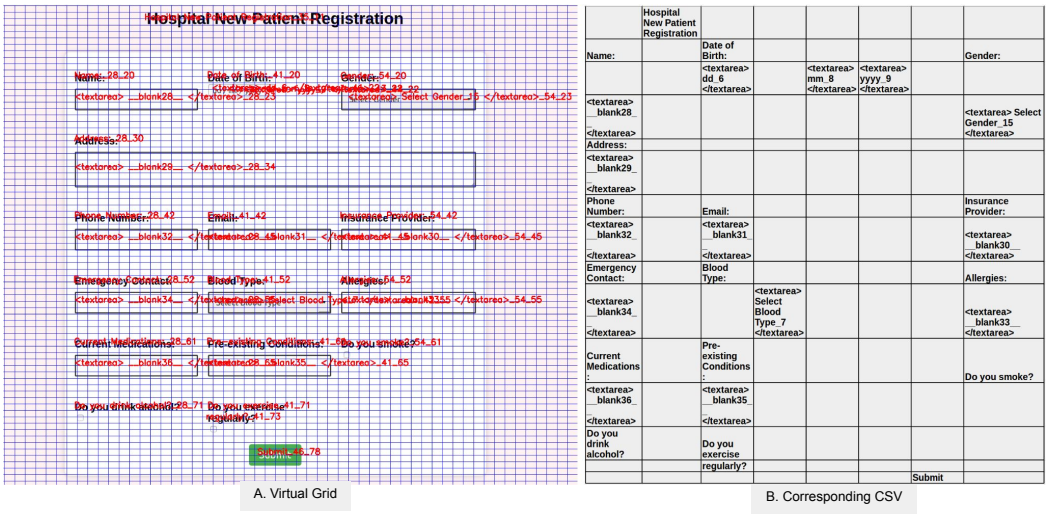

Figure 5: Figure illustrating Layout Mapping using Virtual Grid approach. This approach condenses the original layout by converting pixel coordinates into a virtual grid space. The figure also shows the spatial layout using the .CSV format in virtual grids.

Table 1: Table showing the performance of SmartFlow on Patient Registration application.

Layout No.	Page No.	Accuracy						Task completion average time (in mins)	Complex Component Accuracy		
		OCR		Layout Mapping		Filled Data	Request Submission		Datepicker	Dropdown	Radio/Checkbox
		CER	WER	Rule-based	Virtual-Grid						
1	1	0.0	0.0	0.8	1.0	0.96	1.0	1.99	1.0	1.0	0.867
2	1	0.0	0.0	1.0	0.733	0.91	1.0	1.59	1.0	1.0	0.9
3	1	0.0	0.0	0.8	0.812	0.96	1.0	2.29	1.0	1.0	0.8
4	1	0.0	0.0	1.0	0.833	0.96	1.0	2.16	0.8	1.0	0.93
5	1	0.0	0.0	1.0	1.0	0.97	1.0	1.73	1.0	1.0	1.0
Average		**0.0**	**0.0**	**0.92**	**0.876**	**0.952**	**1.0**	**1.952**	**0.96**	**1.0**	**0.88**

Table 2: Table showing the performance of SmartFlow on Sales Lead Generation application.

Layout No.	Page No.	Accuracy						Task completion average time (in mins)	Complex Component Accuracy		
		OCR		Layout Mapping		Filled Data	Request Submission		Datepicker	Dropdown	Radio/Checkbox
		CER	WER	Rule-based	Virtual-Grid						
1	1	0.044	0.083	0.727	0.636	0.885	1.0	1.37	-	1.0	0.6
2	1	0.0	0.0	1.0	0.833	0.885	1.0	1.38	-	1.0	0.5
3	1	0.003	0.016	0.875	1.0	0.925	1.0	1.71	-	1.0	0.7
4	1	0.0	0.0	1.0	1.0	0.9	1.0	1.12	-	1.0	0.4
5	1	0.031	0.099	1.0	0.733	0.844	1.0	2.2	-	1.0	0.3
Average		**0.015**	**0.039**	**0.92**	**0.841**	**0.887**	**1.0**	**1.55**	**-**	**1.0**	**0.5**

Table 3: Table showing the performance of SmartFlow on Customer Complaint application.

Layout No.	Page No.	Accuracy						Task completion average time (in mins)	Complex Component Accuracy		
		OCR		Layout Mapping		Filled Data	Request Submission		Datepicker	Dropdown	Radio/Checkbox
		CER	WER	Rule-based	Virtual-Grid						
1	1	0.010	0.033	0.9	1.0	0.771	1.0	1.19	-	1.0	0.5
2	1	0.011	0.035	0.92	1.0	0.96	1.0	1.40	1.0	1.0	1.0
3	1	0.011	0.049	1.0	1.0	0.977	1.0	1.43	-	1.0	1.0
4	1	0.004	0.012	1.0	1.0	1.0	1.0	1.19	-	1.0	1.0
5	1	0.005	0.016	1.0	1.0	0.86	1.0	1.59	-	1.0	0.866
Average		**0.008**	**0.029**	**0.964**	**1.0**	**0.913**	**1.0**	**1.36**	**1.0**	**1.0**	**0.873**

Table 4: Table showing the performance of SmartFlow on Passport Registration application.

Layout No.	Page No.	Accuracy						Task completion average time (in mins)	Complex Component Accuracy		
		OCR		Layout Mapping		Filled Data	Request Submission		Datepicker	Dropdown	Radio/Checkbox
		CER	WER	Rule-based	Virtual-Grid						
1	1	0.005	0.019	0.9	1.0	0.94	1.0	1.51	0.8	1.0	0.8
2	1	0.013	0.062	1.0	1.0	0.957	1.0	1.82	1.0	1.0	0.8
3	1	0.007	0.030	0.93	0.93	1.0	1.0	1.71	1.0	1.0	1.0
4	1	0.010	0.046	1.0	1.0	0.92	1.0	1.58	1.0	1.0	0.7
5	1	0.011	0.034	0.81	1.0	1.0	1.0	1.4	1.0	1.0	1.0
Average		**0.009**	**0.038**	**0.928**	**0.986**	**0.963**	**1.0**	**1.604**	**0.96**	**1.0**	**0.86**