

# Supplementary Material for SmartFlow: Robotic Process Automation using LLMs

Arushi Jain, Shubham Paliwal, Monika Sharma, Lovekesh Vig, Gautam Shroff  
TCS Research  
New Delhi, India  
(j.arushi,shubham.p3,monika.sharma1,lovekesh.vig,gautam.shroff)@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 email (CIKM'23)*. ACM, New York, NY, USA, 4 pages. <https://doi.org/XXXXXXX.XXXXXXX>

## 1 ADDITIONAL FIGURES

Due to the 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 displays four examples of diverse application layouts:

- (a) Bank Account Opening Form: Includes fields for Personal Information (First Name, Middle Name, Last Name, Gender, Date of Birth, Address Number), Contact Information (Address, Country, Postal Code, Mobile Number, Email id), and Account Type (Savings, Salary, Certificate of Deposit).
- (b) Patient Registration Form: Includes fields for Name, Date of Birth, Gender, Address, Phone Number, Email, Insurance Provider, Emergency Contact, Blood Type, Allergies, Current Medications, Pre-existing Conditions, and Do you smoke?.
- (c) Conference Attendance System: Includes fields for Conference Details (Conference Name, Edition, Conference Category, Start Date, End Date, Last Registration Date, Participation Type, Country, Location) and International Travel.
- (d) Insurance Claim Processing: Includes fields for Full Name, Email Address, Policy Number, Claim Amount, Description of Claim, and Claim Date.

Figure 1: Figure showing examples of diverse application layouts.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

CIKM'23, Oct 21–25, 2023, Birmingham, UK  
© 2023 Association for Computing Machinery.  
ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00  
<https://doi.org/XXXXXXX.XXXXXXX>

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 4 and Figure 5.

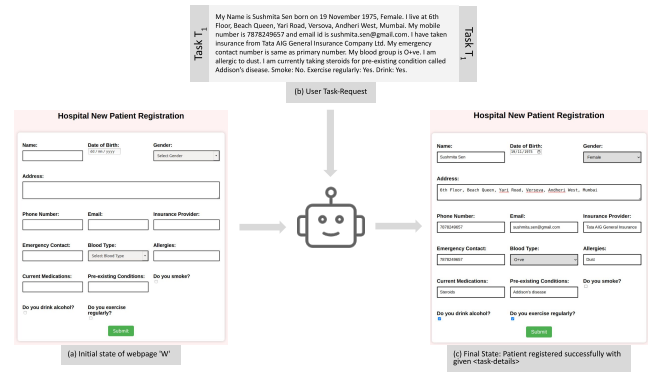
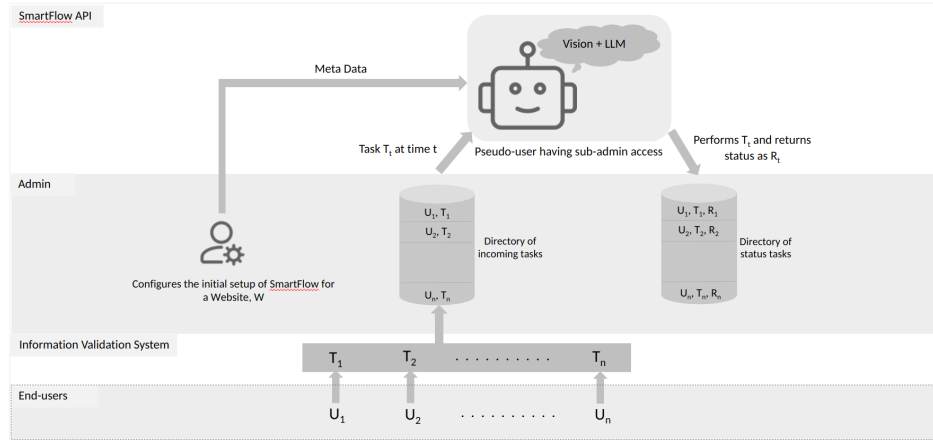


Figure 2: Given a user task-request to register a new patient on the hospital website  $W$  with given details as  $T_1$ , SmartFlow performs the task automatically and returns the status as "Patient registered successfully".

## 2 HANDLING COMPLEX COMPONENTS

Smartflow, an AI-based RPA system, has been designed to handle complex field types such as date pickers, dropdown menus, 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:

- **Datpickers:** They are characterized by their diverse and captivating designs, pose a unique challenge in the realm of user interaction. 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 harnesses the power of vision and a simple technique of the



**Figure 3: The proposed system overview of SmartFlow RPA:** The process begins with end users ( $U_1, U_2, \dots, U_n$ ) submitting corresponding task-requests ( $T_1, T_2, \dots, T_n$ ) 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.

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

Name: <input type="text" value="John"/>	Date of Birth: <input type="text" value="12/20/2003"/> (or Birth)	Gender: <input type="text" value="Select Gender"/>
Address: <input type="text" value=""/>		
Phone Number: <input type="text" value=""/>	Email: <input type="text" value=""/>	Insurance Provider: <input type="text" value=""/>
Emergency Contact: <input type="text" value=""/> Emergency Contact:	Blood Type: <input type="text" value="Select Blood Type"/> Blood Type:	Allergies: <input type="text" value=""/> Allergies:
Current Medications: <input type="text" value=""/> Current Medications:	Pre-existing Conditions: <input type="text" value=""/> Pre-existing Conditions:	Do you smoke? <input type="text" value=""/>
Do you drink alcohol? <input type="text" value=""/>	Do you exercise regularly? <input type="text" value=""/>	

### C. Output Layout Mapping File

## Hospital New Patient Registration

Name: 28\_20

<textarea> \_\_blank28\_ </textarea>\_28\_23

Date of Birth: 41\_20

<input type="text" value="dd 6"/> <input type="text" value="mm 8"/> <input type="text" value="yyyy 9"/>

Gender: 54\_20

<input type="text" value="Select Gender\_15"/>\_54\_25

Address: 28\_30

<textarea> \_\_blank29\_ </textarea>\_28\_34

Phone Number: 28\_42

<input type="text" value="\_\_blank32\_"/>\_28\_45

Email: 41\_42

<input type="text" value="\_\_blank31\_"/>\_41\_45

Insurance Provider: 54\_42

<input type="text" value="\_\_blank30\_"/>\_54\_45

Emergency Contact: 28\_52

<input type="text" value="\_\_blank34\_"/>\_28\_55

Blood Type: 41\_52

<input type="text" value="Select Blood Type\_16"/>\_41\_55

Allergies: 54\_52

<input type="text" value="\_\_blank33\_"/>\_54\_55

Current Medications: 28\_61

<input type="text" value="\_\_blank36\_"/>\_28\_65

Pre-existing Conditions: 41\_60

<input type="text" value="\_\_blank35\_"/>\_41\_65

Do you drink alcohol? 28\_71

<input type="text" value="\_\_blank37\_"/>\_28\_75

Do you exercise regularly? 41\_73

<input type="text" value="\_\_blank38\_"/>\_41\_75

Submit: 28\_78

Hospital New Patient Registration					
Name:		Date of Birth:			Gender:
<input type="text" value="blank28"/>		<input type="text" value="dd 6"/>	<input type="text" value="mm 8"/>	<input type="text" value="yyyy 9"/>	
<input type="text" value="Address: __blank29_"/>					<input type="text" value="Select Gender_15"/>
Phone Number:		Email:			Insurance Provider:
<input type="text" value="blank32_"/>		<input type="text" value="blank31_"/>			<input type="text" value="blank30_"/>
Emergency Contact:		Blood Type:			Allergies:
<input type="text" value="blank34_"/>		<input type="text" value="Select Blood Type_16"/>			<input type="text" value="blank33_"/>
Current Medications:		Pre-existing Conditions:			Do you smoke?
<input type="text" value="blank36_"/>		<input type="text" value="blank35_"/>			
Do you drink alcohol?		Do you exercise regularly?			
					Submit

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. This CSV file is then fed into LLM as input with a prompt to give the mapping of field names with the corresponding edit-fields and data-hints.

**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