

PROAGENT: FROM ROBOTIC PROCESS AUTOMATION TO AGENTIC PROCESS AUTOMATION

Yining Ye^{1*}, Xin Cong^{1*†}, Shizuo Tian¹, Jiannan Cao², Hao Wang³, Yujia Qin¹,
Yaxi Lu¹, Heyang Yu¹, Huadong Wang⁴, Yankai Lin⁵, Zhiyuan Liu^{1†}, Maosong Sun¹

¹Tsinghua University ²Massachusetts Institute of Technology

³Carnegie Mellon University ⁴ModelBest Inc. ⁵Renmin University of China

yeyn2001@gmail.com, xin.cong@outlook.com
tsz21@mails.tsinghua.edu.cn, jiannan@mit.edu, hwang4@alumni.cmu.edu

ABSTRACT

From ancient water wheels to robotic process automation (RPA), automation technology has evolved throughout history to liberate human beings from arduous tasks. Yet, RPA struggles with tasks needing human-like intelligence, especially in elaborate design of workflow construction and dynamic decision-making in workflow execution. As Large Language Models (LLMs) have emerged human-like intelligence, this paper introduces AGENTIC PROCESS AUTOMATION (APA), a groundbreaking automation paradigm using LLM-based agents for advanced automation by offloading the human labor to agents associated with construction and execution. We then instantiate PROAGENT, an LLM-based agent designed to craft workflows from human instructions and make intricate decisions by coordinating specialized agents. Empirical experiments are conducted to detail its construction and execution procedure of workflow, showcasing the feasibility of APA, unveiling the possibility of a new paradigm of automation driven by agents. Our code is public at <https://github.com/OpenBMB/ProAgent>.

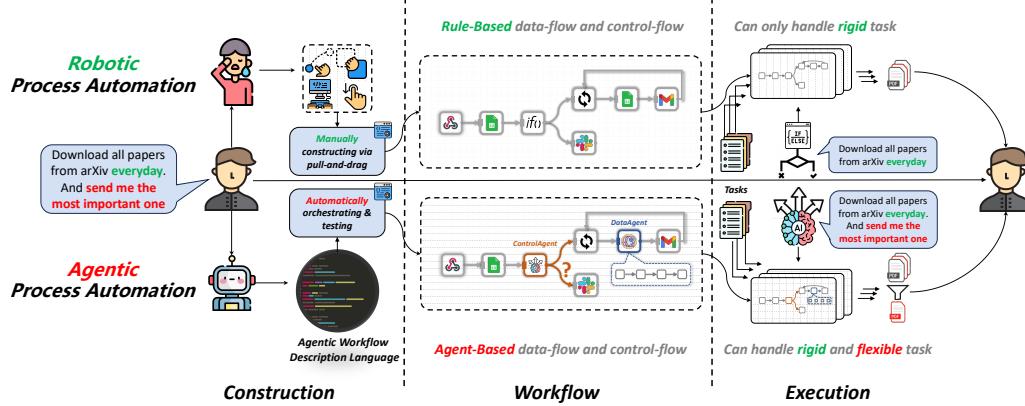


Figure 1: The comparison between Robotic Process Automation and Agentic Process Automation.

1 INTRODUCTION

Automation, aiming to reduce human intervention in processes and enhance efficiency, has undergone a series of evolutionary stages throughout history. From the waterwheel irrigation system in the early agricultural age to steam engines in the industrial age, the human race has continuously been pursuing to offload human labor to autonomous systems, liberating themselves from arduous processes. Entering the information age, marked by a rapid shift from traditional industry to

* Indicates equal contribution.

† Corresponding author.

Paradigm	Efficiency		Intelligence	
	Data Flow	Control Flow	Data Flow	Control Flow
RPA	✓	✓	✗	✗
LLM-based Agents	✗	✗	✓	✓
APA	✓	✓	✓	✓
DataAgent	✓	✓	✓	✗
ControlAgent	✓	✓	✗	✓

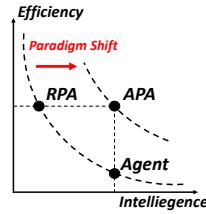


Table 1: A comparison between robotic process automation and agentic process automation in terms of efficiency and flexibility.

an economy primarily based on digital technology, software has been widely used as it serves as the foundation for the processing, storage, and communication of information. Robotic Process Automation (RPA) (Ivančić et al., 2019; Wewerka & Reichert, 2020; Agostinelli et al., 2020; Ferreira et al., 2020)), the *de facto* predominant automation technology, thus has been widely applied, which automates a process by orchestrating several software by manual-crafted rules into a solidified workflow for efficient execution (Zapier; n8n; unipath). Despite its strides, **robotic process automation merely offloads simple and mechanical human labor, while processes requiring human intelligence still necessitate human labor.** First, as Figure 1 shows, while RPA workflows can perform processes automatically, their construction still requires human intelligence for elaborate design. Second, many tasks performed by humans are characterized by their flexible and complex nature while workflows are limited to mechanistically replicating human behavioral processes, posing challenges in automating intricate processes that demand dynamic decision-making capabilities during execution.

With the rapid development of Large Language Models (LLMs) (OpenAI, 2022; 2023), LLMs are emerging with intelligence that was previously exclusive to human beings (Wei et al., 2022). Recently, LLM-based agents have garnered significant attention from the research community (Xi et al., 2023; Wang et al., 2023b; Yao et al., 2022b; Shinn et al., 2023; Sumers et al., 2023; Qin et al., 2023c; Ye et al., 2023). LLM-based agents have demonstrated a certain level of human intelligence, being capable of using tools (Schick et al., 2023; Qin et al., 2023b;c), creating tools (Qian et al., 2023b; Cai et al., 2023), playing games (Wang et al., 2023a; Chen et al., 2023), browsing website (Nakano et al., 2021; Qin et al., 2023a; Yao et al., 2022a), developing software (Qian et al., 2023a) akin to humans. Consequently, a meaningful inquiry naturally emerges: **Can LLM-based agents advance automation in processes necessitating human intelligence, further liberating human beings?**

In this paper, we propose AGENTIC PROCESS AUTOMATION (APA), a novel process automation paradigm that overcomes the two aforementioned limitations of automation. (1) Agentic Workflow Construction: Upon receiving human requirements or instructions, LLM-based agents elaborately construct the corresponding workflows instead of humans. If a process involves dynamic decision-making, agents should recognize which part of this process needs the dynamic decision-making and then orchestrate agents into the workflow. (2) Agentic Workflow Execution: Workflows should be monitored by agents and once the workflow is executed in the dynamic part, agents would intervene to handle the dynamic decision-making.

To explore the feasibility of APA, we instantiate PROAGENT, an LLM-based agent that integrates the agentic workflow construction and agentic workflow execution in a unified framework. For agentic workflow construction, to make LLM-based agents understand and generate workflows, we design **Agentic Workflow Description Language** based on the JSON structure and Python code, stemming from the realization that LLMs are pretrained on coding corpus. Specifically, it adopts JSON structure to organize the input and output data for each software for data standardization and uses Python code to implement process control logic to orchestrate software (see in Figure 2). Upon receiving a specific task, PROAGENT is able to generate the corresponding workflow language to facilitate the construction of the requisite workflow. For agentic workflow execution, dynamic decision-making in workflows encompasses two aspects: (1) Data flow: complex data processing (e.g., writing data analysis reports) often exceed the capacity of rule-based systems and thus agents must intervene to effectively manage these intricate processes. (2) Control flow: complex tasks may involve intricate conditional branches and loops, which surpass the expression ability of rules. In such cases, agents need to function as controllers to dynamically determine the subsequent actions. Hence, we

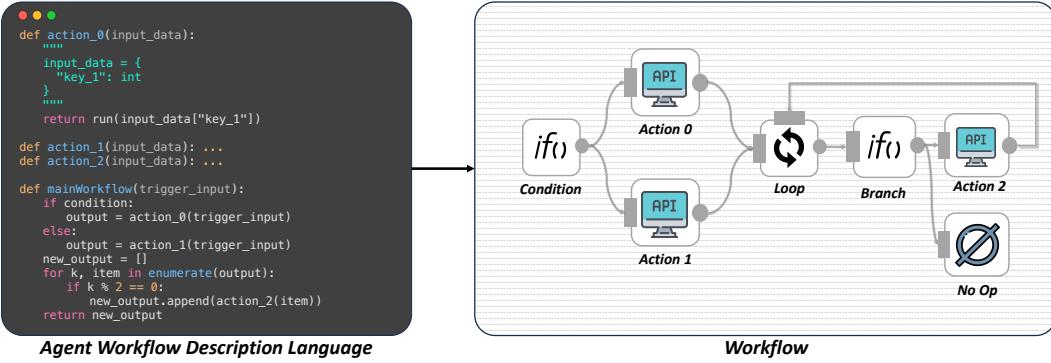


Figure 2: Illustration of Agentic Workflow Description Language.

design two types of dynamic decision-making agents: **DataAgent** acts as a data processing to handle intricate data processes dynamically and **ControlAgent** functions as a condition expression that enables the dynamic determination of subsequent branches for execution. Confronted with complex tasks that need intelligence, PROAGENT can orchestrate these two agents into the workflows during construction and handle complex circumstances purposefully during execution, offloading the intelligent labor (see in Table 1).

To empirically validate our approach, we conduct proof-of-concept experiments to showcase that PROAGENT is able to construct workflows based on human instructions and handle the dynamic decision-making part of the process by utilizing agents in the workflow. We further discuss the relationship between PROAGENT with existing research areas, including Tool Learning (Qin et al., 2023b;c), Process Mining (Tiwari et al., 2008; Van Der Aalst, 2012; Turner et al., 2012), Safety (Cummings, 2004) and etc. Our contributions are listed as follows:

- We propose AGENTIC PROCESS AUTOMATION, a new process automation paradigm that integrates LLM-based agents to further offload the intelligent labor of humans.
- We instantiate PROAGENT, in which Agentic Workflow Description Language is designed for LLM-based agents to construct workflows and DataAgent and ControlAgent are orchestrated into workflows to handle the dynamic decision-making process part purposefully.
- We demonstrate of feasibility of our PROAGENT through proof-of-concept case analyses and the exploration of potential and opportunities of AGENTIC PROCESS AUTOMATION across various research domains including tool learning, process mining, safety, etc.

2 METHODOLOGY

Workflow is widely-used in RPA to solidify the process by a software invocation graph, where nodes represent a software operation and edges signify topology of the process of execution. To achieve the solidification, a data flow and a control flow are involved to within the workflow. Data flow describes how data is passed and processed within a series of software and control flow describes the order of software to execute. In this section, we first introduce Agentic Workflow Description Language to express the data flow and control flow, and then we further detail how to integrate agents into workflows to bring flexibility into workflows. Finally, we detail the workflow construction and execution procedure about how PROAGENT works.

2.1 AGENTIC WORKFLOW DESCRIPTION LANGUAGE

As workflow is a graph-based representation approach for RPA to solidify the process, it is inadaptive to LLMs to understand and generate workflows. Thus, we elaborate design Agentic Workflow Description Language for LLM-agents to conveniently solidify workflows based on the characteristics of coding pretraining. Specifically, we adopt JSON structure to describe data flow

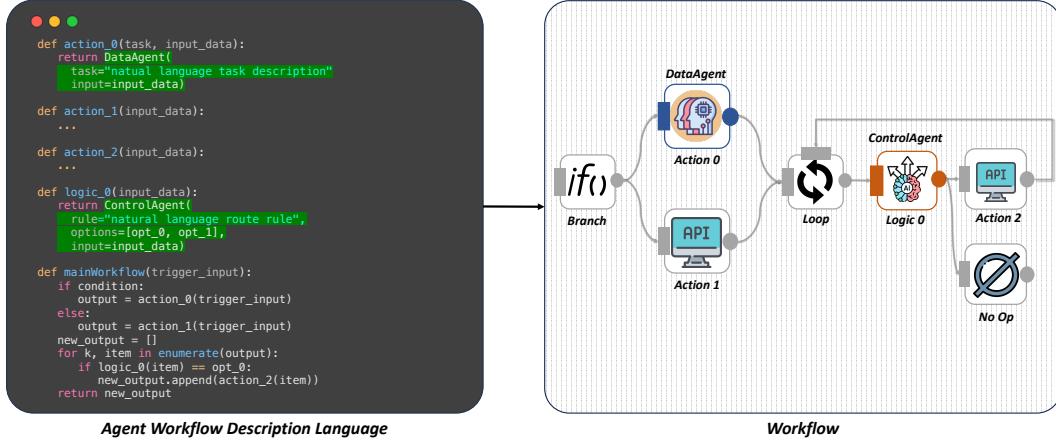


Figure 3: Illustration of Agentic Workflow Description Language with DataAgent and ControlAgent.

and Python code to describe control flow. Figure 2 gives the illustration of Agentic Workflow Description Language.

JSON Structure for Data Flow To solidify a workflow, the data format through software should be standardized to ensure the automatic data process, free from unnecessary agent interventions. We adapt the JSON structure to organize the input/output data of all actions in the workflow. As Figure 2 shows, the input data is formatted in a key-value-paired dictionary. Every data should be assigned a specific key, making it easy to parse and manipulate. When transferring data between different software, the JSON structure is convenient to index the specific data field. Only when the input and output of all software are strictly standardized, promoting consistency across different software of the workflow, thereby reducing the likelihood of data interpretation errors or discrepancies.

Python Code for Control Flow For complex tasks, the corresponding workflows usually involve complex control logic, including conditional branches, loops, or sub-workflow execution. Conventional RPA methods commonly design graph-based representations for human developers to describe the control flow (Zapier; n8n; unipath) but its expression ability for complex workflow is limited and it is also not suitable for LLM-based agents to understand and generate. As Python programming language supports complex control logic and more importantly and it is learned by LLMs during the pre-training phase, we use Python to describe the control flow. As a high-level programming language, Python offers a rich set of primitives and features, providing greater expressive capability to describe complex control logic. A workflow is composed of a Python file, with each software operation aligned to a Python function called *action*. The corresponding input/output data is mapped into the parameters and return values of the function. Thus, a series of actions (i.e., software) are described as sequential function callings in Python. The if-else statement and for/while statement in Python can be used to implement complex logic control flow. Finally, the workflow is encapsulated within a main Python function (i.e., `mainWorkflow`). Furthermore, as Python supports the nested function calling, different workflows can also be composed together by calling workflow function to construct a complex workflow. During workflow execution, we utilize a Python executor, starting from the main workflow function (`mainWorkflow`) as the entry point and execute each functions sequentially, ultimately completing the entire workflow execution.

2.2 AGENT-INTEGRATED WORKFLOW

As many real-world tasks with flexibility and complexity nature involve dynamic decision-making process, we devise DataAgent and ControlAgent which can be orchestrated into workflows to handle the dynamic part during execution. Figure 3 gives the illustration.

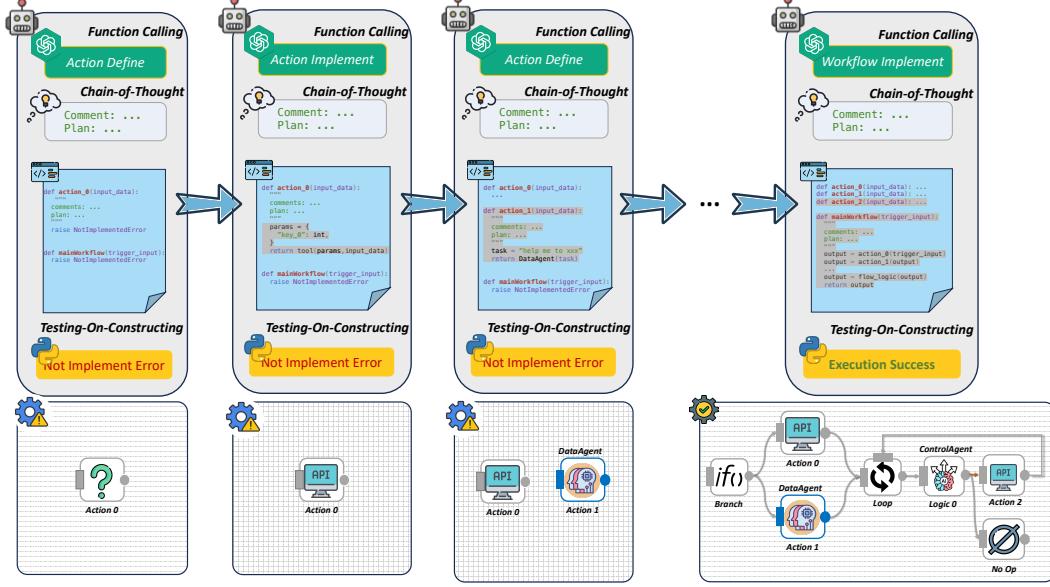


Figure 4: The Illustration of the workflow construction procedure of PROAGENT.

DataAgent To achieve complex data process, we devise DataAgent, which acts as an action that is operated by an LLM-based agent. As Figure 3 shows, it supports inputting a task description based on natural language and then accomplishing this task autonomously based on the intelligence of the agent. During execution, this function initiates a ReACT-based agent (Yao et al., 2022b) to fulfill the task.

$$\text{output} \leftarrow \text{DataAgent}(\text{task}, \text{input}) \quad (1)$$

Although the function is actually operated by agents, its input/output data are still organized by JSON to make it can be orchestrated into existing workflows to connect with other actions. By incorporating the DataAgent, the workflow provides support for enhanced flexibility for data flow, enabling the handling of intricate data processing demands.

ControlAgent In addition to serving as the action, agents can be further involved in the control flow to schedule the execution logic. We introduce ControlAgent into the control flow, allowing it to substitute a selection expression. As Figure 3 shows, ControlAgent contains a pre-generated judgment criterion based on natural language and several execution branch candidates.

$$\text{opt} \leftarrow \text{ControlAgent}(\text{task}, \text{input}, [\text{opt}_1, \text{opt}_2, \dots, \text{opt}_n]) \quad (2)$$

During execution, the agent can make a decision based on the input data to decide which branch will be executed subsequently, influencing the control flow of the workflow.

2.3 WORKFLOW CONSTRUCTION

As the workflow is represented as JSON structure and Python code, the workflow construction is formulated as a code generation task. As Figure 4 demonstrates, the workflow construction procedure contains four iterative operations:

- `action_define`: It determines which action is selected to add into the workflow.
- `action_implement`: It first transforms the action into the Python function by determining its input/output data format in JSON structure and then implement the data process program in Python code.
- `workflow_implement`: As workflows are represented as `mainWorkflow` functions, this operation refers to providing an implementation for it to orchestrate the entire workflow.
- `task_submit`: It is used to denote the termination of the workflow construction.

In practice, we employ OpenAI GPT-4 as the backbone of PROAGENT to generate the workflow language and further incorporated several techniques to enhance the workflow generation capabilities:

- **Testing-on-Constructing:** During the construction, PROAGENT tends to test each function or entire workflow, which ensures the validation of the constructed workflow before execution.
- **Function Calling:** The aforementioned four operations are defined as function in GPT-4 to use Function Calling to explicitly control the whole construction procedure, benefiting controllable generation.
- **Chain-of-Thought:** When implementing each function, PROAGENT requires to provide a comment (explaining the purpose of this function) and a plan (indicating what the subsequent operations should be done next), which aids in enhancing the workflow code generation performance.

2.4 WORKFLOW EXECUTION

The workflow execution procedure is based on Python interpreter. Given a workflow language, once this workflow is triggered, its corresponding `mainWorkflow` function is selected as the entry point to begin the execution procedure. The execution procedure follows the Python code execution rule, i.e., executing according to the line order sequentially. Once the `mainWorkflow` function returns, the workflow execution is finished successfully.

3 PROOF-OF-CONCEPT EXPERIMENT

To validate the feasibility of AGENTIC PROCESS AUTOMATION, we conduct proof-of-concept experiment based on n8n¹, an open-source workflow platform. Each APP (i.g., software) in the n8n platform is encapsulated as an action in the workflow and thus the core of the workflow construction is to orchestrate these APPs to achieve certain tasks. We implement our proposed PROAGENT based on GPT-4. We construct a case about the commercial scenario to explain how our PROAGENT works in detail.

3.1 TASK CONSTRUCTION

Task
<p>When I send a worksheet of business lines through Web, deal with them according to which type of each business line belong to.</p> <ol style="list-style-type: none"> 1. To-Customer: Send a message to Slack to report the profits of business lines. 2. To-Business: Write a report which should analyze the data to give some suggestions and then send it to the Gmail of the corresponding managers.

We present a typical commercial scenario where a business department manager seeks to extract diverse business line data from Google Sheets. Upon identifying a business line is To-Customer, a message is dispatched to the Slack channel. In the case of a business line as To-Business, an email is sent to the respective manager, encompassing an assessment of the business line and a concise overview of its profitability. This exemplifies the capabilities achievable through the integration of workflow and agent:

- Given the similar format of the content in Google Sheets, this task exhibits a high level of reusability, applicable to various business departments with analogous requirements.
- Distinguishing whether a business line pertains To-Customer or To-Business based solely on its description is a complex task not easily handled through rule-based judgments. Agent integration becomes imperative for dynamic branching.
- Generating corresponding emails for the relevant business lines is a dynamic requirement necessitating flexible agent-based data processing.

¹<https://n8n.io>

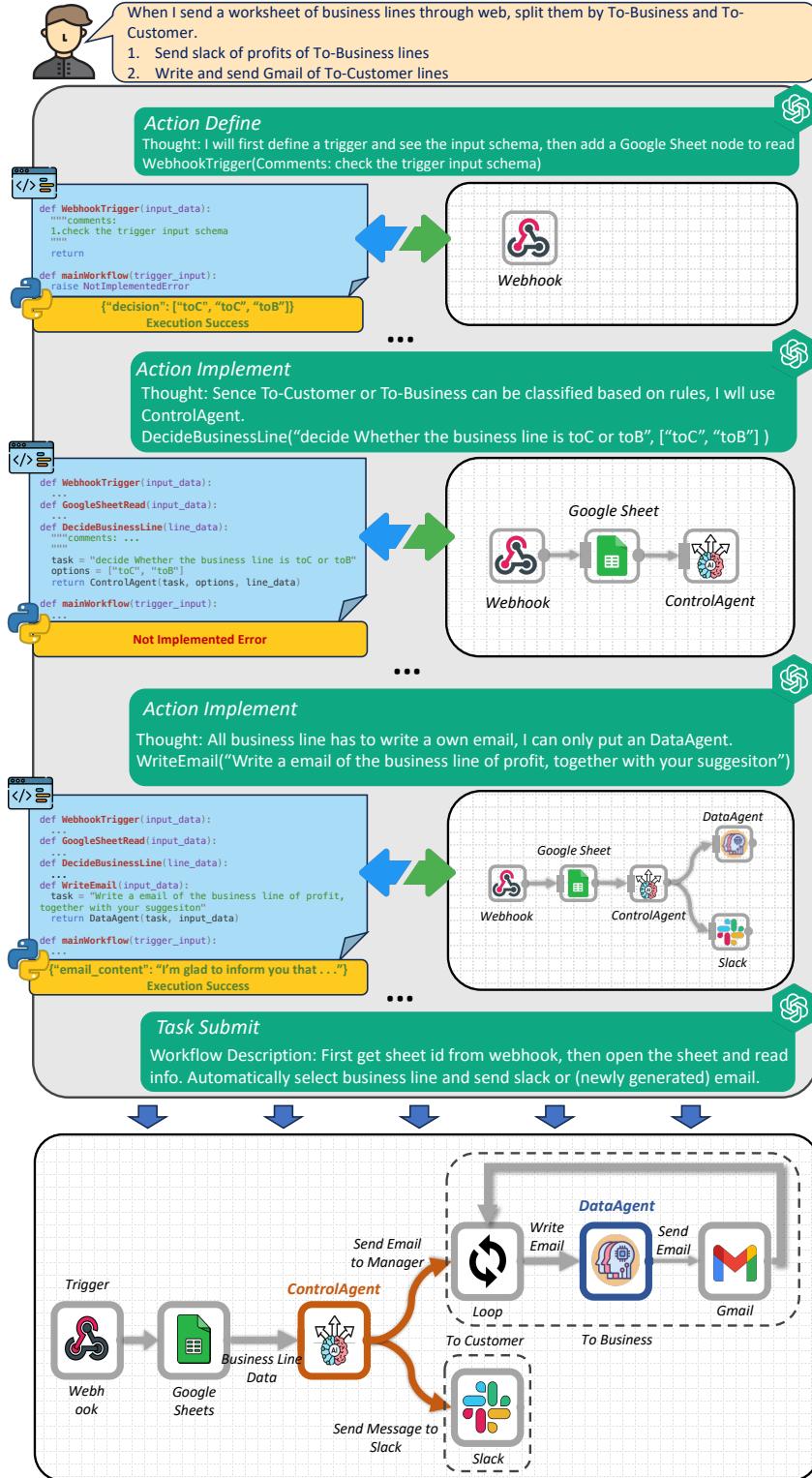


Figure 5: The Illustration of the workflow construction procedure of PROAGENT for case analysis.

3.2 WORKFLOW CONSTRUCTION

Figure 5 gives the visualization of the constructed workflow by PROAGENT. PROAGENT constructs a workflow with seven nodes for this case, including a DataAgent node and a ControlAgent node.

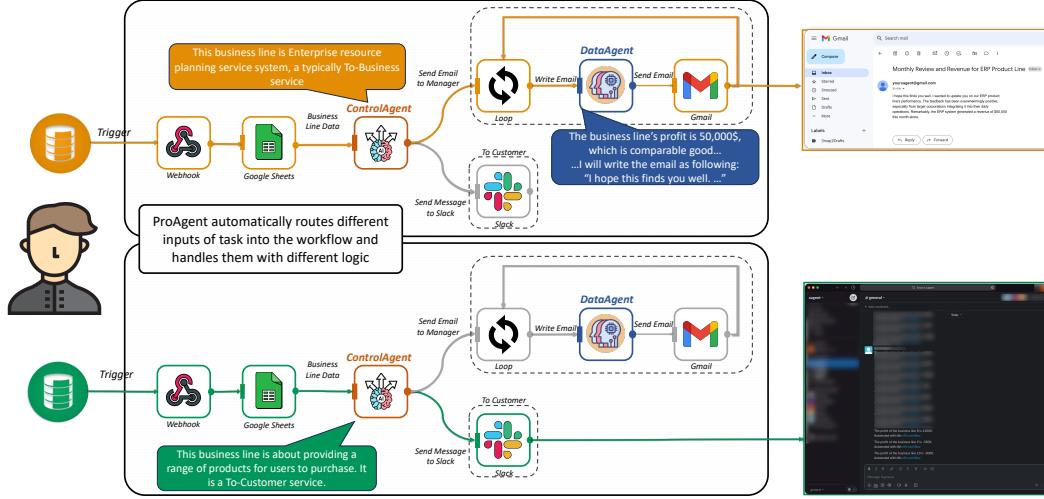


Figure 6: The Illustration of the workflow execution procedure of PROAGENT for case analysis.

As the user says that he will send data through Web, PROAGENT decide to define an action named `WebhookTrigger` as the trigger to for the workflow. Then it implement `GoogleSheetRead` action to read data from Google Sheets according to the user description. Since it should execute different actions according to whether the business line belong to To-Business or To-Customer which needs to understand the meaning of each business line, PROAGENT define a `ControlAgent` which aims to *decide whether the business line is toC or toB* to judge what next action to execute. If the business line belong to To-Customer, as the user description, PROAGENT implements the `Slack` action which send the corresponding profits into Slack. If the business line belong to To-Business, it needs to write a detailed report to analyze the specific data in Google Sheets and give some suggestions. Thus, PROAGENT implement a `DataAgent WriteEmail` which task is to *Write a email of the business line of profit, together with your suggestions*. Then, a Gamil APP is following implemented to send the generated email to the corresponding managers. As there may exist multiple To-Business data in Google Sheets, PROAGENT further add a `Loop` in workflow to deal with these data iteratively. Finally, as the workflow is constructed completely, `task_submit` is operated by PROAGENT to end the construction procedure.

3.3 WORKFLOW EXECUTION

Figure 6 illustrates two execution cases for the constructed workflow. These two cases demonstrate a To-Customer and a To-Business line respectively. It is obviously shown that the `ControlAgent` successfully distinguish which type of two business lines belong to. For the first one, the description of this business line is:

“Enterprise Resource Planning: organizations use to manage day-to-day business activities such as accounting, procurement, project management, and risk management, and supply chain operations.”

`ControlAgent` distinguish that *This business line is Enterprise resource planning system, a typically to-business service*. After that, the `DataAgent` is executed to write an email to send:

“I hope this finds you well. I wanted to update you on our ERP product line’s performance. The feedback has been overwhelmingly positive, especially from larger corporations integrating it into their daily operations. Remarkably, the ERP system generated a revenue of \$50,000 this month alone.”

For the second one, its description is:

“E-commerce Marketplace: Operating an online platform for consumers to purchase a wide range of products from various brands and sellers.”

PROAGENT extracts some key words such as “products” and “purchase” to decide that this business line belongs to To-Customer type. Then, a simple message “*The profit of the business line 9 is -3500\$.*” to the Slack.

4 DISCUSSION

In this section, we discuss the relationship between APA and PROAGENT technologies and the field of tool learning, process mining, etc, along with potential research directions. Subsequently, we discuss the issue of automation bias that may arise from APA technologies and delve into the unique value humans should bring in the era of automation.

4.1 PROAGENT AS TOOLAGENT: INTEGRATING TOOL UTILIZATION AND TOOL CREATION

Existing work has shown that LLM-based agents own powerful capability to utilize external tools (e.g., search engine, web browser, calculator, python interpreter, etc) (Ivančić et al., 2019; Wewerka & Reichert, 2020; Agostinelli et al., 2020; Ferreira et al., 2020). By integrating the external tools, agents can extend its capability from text generation to manipulate tools to impact environments and even the real world. In addition to their adeptness in utilizing existing tools, recent research has showcased their capability to create novel tools for specific tasks (Qian et al., 2023b; Cai et al., 2023). In a workflow, a series of operations are involved, and when these operations involve invoking external tools, the execution process of this workflow can be viewed as the process of tool utilization (Nakano et al., 2021; Yao et al., 2022a; Schick et al., 2023; Qin et al., 2023b;c; Ye et al., 2023). In the context of workflow management, the execution of a series of operations through the invocation of external tools can be perceived as a process of tool utilization. Orchestrating multiple operations (i.e., tools) into a coherent workflow to address specific tasks can be viewed as the process of tool creation, i.e., combining multiple tools to form a new tool. In this context, PROAGENT can be understood as an autonomous process that seamlessly integrates tool creation and utilization.

4.2 PROCESS MINING: DISCOVER, ANALYZE, AND IMPROVE WORKFLOWS

In our PROAGENT, the workflow construction process is primarily driven by the agent solely, potentially yielding sub-optimal workflows. Integrating the principles of Process Mining (PM) (Tiwari et al., 2008; Van Der Aalst, 2012; Turner et al., 2012) into this framework can facilitate the discovery of valuable workflows from historical records. Process Mining is a data-driven technique used to discover, analyze, and improve existing workflows. It involves the extraction of insights and knowledge from historical records generated during the execution of various workflows. This method enables to gain a comprehensive understanding of workflows constructed by agents, identify inefficiencies, bottlenecks, and deviations from the intended workflow, and subsequently optimize their operations for better performance and efficiency. Moreover, optimized workflows can serve as effective training data to enhance the agent’s workflow construction capabilities.

4.3 ETHICAL AND SAFETY CONCERNs: AGENT AGGRAVATES AUTOMATION BIAS

The advent of agent technology offers the potential for models to tackle complex tasks. However, on the flip side, it can also lead to an over-reliance on agents by humans. People might opt to trust agent decisions even when these decisions conflict with their own views, which is known as Automation Bias (Cummings, 2004) and has been observed in domains such as Clinical decision support systems (Goddard et al., 2012). Under LLM-based agent technologies, this problem becomes even more pronounced. Humans may shift their trust in the stability of traditional rule-based workflows to agents, mistakenly believing that the agent’s decision-making processes are equally reliable, especially hallucination (Maynez et al., 2020; Zhang et al., 2023; Ji et al., 2023). Future research is

necessary to prioritize the development of safer, more trustful, more interpretable agentic process automation.

4.4 HUMAN ADVANTAGE: RETHINKING THE MEANING OF HUMAN LABOR

APA introduces the intelligence of elaborate design in workflow construction and dynamic decision-making in workflow execution into process automation, which can offload the heavy human labor in RPA. Now, the more pertinent question is: “*What tasks should remain human-driven?*” There are processes that inherently benefit from human intuition, experience, and creativity. For these tasks, humans play a crucial role that can’t be easily supplanted by machines. While automation might offer efficiency, it can’t replicate the nuanced understanding and innovative solutions that a human brings to the table. The paradox of human involvement, where human intervention can improve outcomes, stands in contrast to the earlier mentioned pitfalls of automation bias.

The next frontier in APA involves discerning which processes can be wholly automated and which require human oversight or intervention. We must remember that the ultimate goal of automation is to amplify productivity, not to supplant humans entirely. The challenge lies in facilitating a symbiotic relationship between humans and machines, where neither is completely excluded in favor of the other. Drawing from the perspective of Steve Jobs, the future should see humans focusing on what they do best: applying their unique intelligence and creativity where it matters most. APA demands a recalibration, where automation serves humanity, and humans, in turn, elevate the capabilities of automation.

5 RELATED WORK

Robotic Process Automation Robotic process automation (RPA) (Ivančić et al., 2019; Hofmann et al., 2020; Tiwari et al., 2008; Scheer et al., 2004), as the fashion automation paradigm, primarily employs software robots to either automate access to software APIs or simulate user GUI interactions to accomplish tasks through multiple software. Unlike traditional automation techniques, RPA emulates the way humans use software, directly tapping into existing software assets without the need for transformation or additional investment. Thus, RPA has gained substantial attention in recent years as an effective technology for automating repetitive and rule-based tasks typically performed by human workers (Zapier; n8n; unipath). RPA are primarily designed to automate repetitive tasks using predefined rules and workflow templates, which needs heavy human labor to design and implement workflows. Still, due to the workflows are driven by manual-crafted rules, it struggles to handle those complex tasks that needs dynamic decision-making.

Recently, there has been a growing interest in integrating RPA with AI technique, leading to various terminologies and definitions. For instance, *Intelligent Process Automation* (IPA) (Ferreira et al., 2020; Chakraborti et al., 2020b) and *Cognitive Automation* (or RPA 4.0) (Lacity & Willcocks, 2018), aim to amalgamate AI techniques in the phases of RPA, e.g., data format transformation (Leno et al., 2020), workflow optimization (Chakraborti et al., 2020a), conversational assistant (Moiseeva et al., 2020), demonstration-to-process translation (Li et al., 2019), etc. Nevertheless, these work still utilizes traditional deep learning technique (e.g., RNN (Han et al., 2020)) or even machine learning technique (e.g., Monte Carlo Tree Search (Chen, 2020)) into RPA. More importantly, they just utilizes AI technique into some specific fragments of RPA (e.g., data format transformation (Leno et al., 2020)). In contrast, our work AGENTIC PROCESS AUTOMATION takes the lead to integrate the most intelligent AI model, large language models, into RPA. As a result, it is the inaugural exploration into agentic techniques in both the autonomous generation of workflows and Agent-driven workflow execution to endow them with intelligence.

LLM-based Agents Large language models (LLMs), as significant milestones of artificial intelligence, unveil the remarkable capability on a wide range of tasks (OpenAI, 2022; 2023). Recently, LLM-based agents emerge to extend LLMs with external tools to interact with the environment to achieve real-world tasks. Early research work attempt to prompt LLMs to generate the action according to the observation of environment (Nakano et al., 2021; Huang et al., 2022; Ahn et al., 2022; Schick et al., 2023; Qian et al., 2023a; Chen et al., 2023). Such a manner tends to struggle when facing intricate tasks that need long-term planning and decision-making. To address this issue, ReAct (Yao et al., 2022b) proposed a dynamic task-solving approach that makes agents generate

thought for each action to form a reasoning chain, enabling flexible reasoning-guided, trackable, and adjustable actions, resulting in notable improvements compared to act-only methodologies. Based on the dynamic task-solving manner, many agents are proposed subsequently to improve agent capability in different aspects, e.g., reflection (Shinn et al., 2023), planning (Yao et al., 2023; Hao et al., 2023; Besta et al., 2023; Sel et al., 2023), tool learning (Schick et al., 2023; Patil et al., 2023; Qin et al., 2023b;c; Qian et al., 2023b), multi-agents (Park et al., 2023; Qian et al., 2023a), etc. However, all the existing ReACT-based agent methods are restricted to linearly generate decision-making, resulting in lower operational efficiency. In this paper, we propose PROAGENT that explores to enhance the efficiency of the dynamic task-solving approach by recognizing which part of the workflow needs the intelligence involves and integrating agents to handle these parts purposefully.

6 CONCLUSION

In this research, we present a novel process automation paradigm, AGENTIC PROCESS AUTOMATION, to address the limitations of robotic process automation technologies in handling tasks requiring human intelligence by harnessing the capabilities of LLM-based agents to integrate them into the workflow construction and execution process. Through the instantiation of PROAGENT, we illustrated how LLM-based agents can feasibly manage complex decision-making processes, thereby offloading the burden of intelligent labor from humans. Our proof-of-concept experiment provided evidence of the feasibility of AGENTIC PROCESS AUTOMATION in achieving efficiency and flexibility in process automation. Our findings contribute to the growing body of research in the field of intelligent automation and underscore the significant role that LLM-based agents can play in enhancing the efficiency and flexibility of various industries. As the adoption of automation technologies continues to expand, we anticipate that the APA framework can serve as a catalyst for further advancements in the automation landscape, leading to increased efficiency, reduced human intervention, and ultimately, a more streamlined and intelligent workflow ecosystem.

REFERENCES

- Simone Agostinelli, Andrea Marrella, and Massimo Mecella. Towards intelligent robotic process automation for bpmers. *arXiv preprint arXiv:2001.00804*, 2020.
- Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David, Chelsea Finn, Keerthana Gopalakrishnan, Karol Hausman, Alex Herzog, et al. Do as i can, not as i say: Grounding language in robotic affordances. *ArXiv preprint, abs/2204.01691*, 2022.
- Maciej Besta, Nils Blach, Ales Kubicek, Robert Gerstenberger, Lukas Gianinazzi, Joanna Gajda, Tomasz Lehmann, Michal Podstawska, Hubert Niewiadomski, Piotr Nyczyk, et al. Graph of thoughts: Solving elaborate problems with large language models. *arXiv preprint arXiv:2308.09687*, 2023.
- Tianle Cai, Xuezhi Wang, Tengyu Ma, Xinyun Chen, and Denny Zhou. Large language models as tool makers. *arXiv preprint arXiv:2305.17126*, 2023.
- Tathagata Chakraborti, Shubham Agarwal, Yasaman Khazaeni, Yara Rizk, and Vatche Isahagian. D3ba: a tool for optimizing business processes using non-deterministic planning. In *Business Process Management Workshops: BPM 2020 International Workshops, Seville, Spain, September 13–18, 2020, Revised Selected Papers 18*, pp. 181–193. Springer, 2020a.
- Tathagata Chakraborti, Vatche Isahagian, Rania Khalaf, Yasaman Khazaeni, Vinod Muthusamy, Yara Rizk, and Merve Unuvar. From robotic process automation to intelligent process automation: –emerging trends–. In *Business Process Management: Blockchain and Robotic Process Automation Forum: BPM 2020 Blockchain and RPA Forum, Seville, Spain, September 13–18, 2020, Proceedings 18*, pp. 215–228. Springer, 2020b.
- Weize Chen, Yusheng Su, Jingwei Zuo, Cheng Yang, Chenfei Yuan, Chen Qian, Chi-Min Chan, Yujia Qin, Yaxi Lu, Ruobing Xie, et al. Agentverse: Facilitating multi-agent collaboration and exploring emergent behaviors in agents. *arXiv preprint arXiv:2308.10848*, 2023.

- Yiru Chen. Monte carlo tree search for generating interactive data analysis interfaces. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, pp. 2837–2839, 2020.
- Mary Cummings. Automation bias in intelligent time critical decision support systems. In *AIAA 1st intelligent systems technical conference*, pp. 6313, 2004.
- Deborah Ferreira, Julia Rozanova, Krishna Dubba, Dell Zhang, and Andre Freitas. On the evaluation of intelligent process automation. *arXiv preprint arXiv:2001.02639*, 2020.
- Kate Goddard, Abdul Roudsari, and Jeremy C Wyatt. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1):121–127, 2012.
- Xue Han, Lianxue Hu, Yabin Dang, Shivali Agarwal, Lijun Mei, Shaochun Li, and Xin Zhou. Automatic business process structure discovery using ordered neurons lstm: a preliminary study. *arXiv preprint arXiv:2001.01243*, 2020.
- Shibo Hao, Yi Gu, Haodi Ma, Joshua Jiahua Hong, Zhen Wang, Daisy Zhe Wang, and Zhiting Hu. Reasoning with language model is planning with world model. *arXiv preprint arXiv:2305.14992*, 2023.
- Peter Hofmann, Caroline Samp, and Nils Urbach. Robotic process automation. *Electronic markets*, 30(1):99–106, 2020.
- Wenlong Huang, Pieter Abbeel, Deepak Pathak, and Igor Mordatch. Language models as zero-shot planners: Extracting actionable knowledge for embodied agents. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 9118–9147. PMLR, 2022.
- Lucija Ivančić, Dalia Suša Vugec, and Vesna Bosilj Vukšić. Robotic process automation: systematic literature review. In *Business Process Management: Blockchain and Central and Eastern Europe Forum: BPM 2019 Blockchain and CEE Forum, Vienna, Austria, September 1–6, 2019, Proceedings 17*, pp. 280–295. Springer, 2019.
- Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12):1–38, 2023.
- Mary Lacity and Leslie P Willcocks. *Robotic process and cognitive automation: the next phase*. SB Publishing, 2018.
- Volodymyr Leno, Marlon Dumas, Marcello La Rosa, Fabrizio Maria Maggi, and Artem Polyvyanyy. Automated discovery of data transformations for robotic process automation. *arXiv preprint arXiv:2001.01007*, 2020.
- Toby Jia-Jun Li, Marissa Radensky, Justin Jia, Kirielle Singarajah, Tom M Mitchell, and Brad A Myers. Interactive task and concept learning from natural language instructions and gui demonstrations. *arXiv preprint arXiv:1909.00031*, 2019.
- Joshua Maynez, Shashi Narayan, Bernd Bohnet, and Ryan McDonald. On faithfulness and factuality in abstractive summarization. *arXiv preprint arXiv:2005.00661*, 2020.
- Alena Moiseeva, Dietrich Trautmann, Michael Heimann, and Hinrich Schütze. Multipurpose intelligent process automation via conversational assistant. *arXiv preprint arXiv:2001.02284*, 2020.
- n8n. n8n.io - a powerful workflow automation tool. URL <https://n8n.io/>.
- Reiichiro Nakano, Jacob Hilton, Suchir Balaji, Jeff Wu, Long Ouyang, Christina Kim, Christopher Hesse, Shantanu Jain, Vineet Kosaraju, William Saunders, et al. Webgpt: Browser-assisted question-answering with human feedback. *ArXiv preprint*, abs/2112.09332, 2021.

- OpenAI. OpenAI: Introducing ChatGPT, 2022. URL <https://openai.com/blog/chatgpt>.
- OpenAI. Gpt-4 technical report, 2023.
- Joon Sung Park, Joseph C O'Brien, Carrie J Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. Generative agents: Interactive simulacra of human behavior. *arXiv preprint arXiv:2304.03442*, 2023.
- Shishir G Patil, Tianjun Zhang, Xin Wang, and Joseph E Gonzalez. Gorilla: Large language model connected with massive apis. *arXiv preprint arXiv:2305.15334*, 2023.
- Chen Qian, Xin Cong, Cheng Yang, Weize Chen, Yusheng Su, Juyuan Xu, Zhiyuan Liu, and Maosong Sun. Communicative agents for software development. *arXiv preprint arXiv:2307.07924*, 2023a.
- Cheng Qian, Chi Han, Yi R Fung, Yujia Qin, Zhiyuan Liu, and Heng Ji. Creator: Disentangling abstract and concrete reasonings of large language models through tool creation. *arXiv preprint arXiv:2305.14318*, 2023b.
- Yujia Qin, Zihan Cai, Dian Jin, Lan Yan, Shihao Liang, Kunlun Zhu, Yankai Lin, Xu Han, Ning Ding, Huadong Wang, et al. Webcpm: Interactive web search for chinese long-form question answering. *arXiv preprint arXiv:2305.06849*, 2023a.
- Yujia Qin, Shengding Hu, Yankai Lin, Weize Chen, Ning Ding, Ganqu Cui, Zheni Zeng, Yufei Huang, Chaojun Xiao, Chi Han, et al. Tool learning with foundation models. *arXiv preprint arXiv:2304.08354*, 2023b.
- Yujia Qin, Shihao Liang, Yining Ye, Kunlun Zhu, Lan Yan, Yaxi Lu, Yankai Lin, Xin Cong, Xiangru Tang, Bill Qian, et al. Toolllm: Facilitating large language models to master 16000+ real-world apis. *arXiv preprint arXiv:2307.16789*, 2023c.
- August-Wilhelm Scheer, Ferri Abolhassan, Wolfram Jost, and Mathias Kirchmer. Business process automation. *ARIS in practice*, 2004.
- Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. Toolformer: Language models can teach themselves to use tools. *ArXiv preprint, abs/2302.04761*, 2023.
- Bilgehan Sel, Ahmad Al-Tawaha, Vanshaj Khattar, Lu Wang, Ruoxi Jia, and Ming Jin. Algorithm of thoughts: Enhancing exploration of ideas in large language models. *arXiv preprint arXiv:2308.10379*, 2023.
- Noah Shinn, Federico Cassano, Beck Labash, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning, 2023.
- Theodore Sumers, Shunyu Yao, Karthik Narasimhan, and Thomas L Griffiths. Cognitive architectures for language agents. *arXiv preprint arXiv:2309.02427*, 2023.
- Ashutosh Tiwari, Chris J Turner, and Basim Majeed. A review of business process mining: state-of-the-art and future trends. *Business Process Management Journal*, 14(1):5–22, 2008.
- Chris J Turner, Ashutosh Tiwari, Richard Olaiya, and Yuchun Xu. Process mining: from theory to practice. *Business Process Management Journal*, 18(3):493–512, 2012.
- unipath. The unipath business automation platform. URL <https://www.unipath.com/>.
- Wil Van Der Aalst. Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems (TMIS)*, 3(2):1–17, 2012.
- Guanzhi Wang, Yuqi Xie, Yunfan Jiang, Ajay Mandlekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, and Anima Anandkumar. Voyager: An open-ended embodied agent with large language models. *arXiv preprint arXiv:2305.16291*, 2023a.

- Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, et al. A survey on large language model based autonomous agents. *arXiv preprint arXiv:2308.11432*, 2023b.
- Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, et al. Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*, 2022.
- Judith Wewerka and Manfred Reichert. Robotic process automation—a systematic literature review and assessment framework. *arXiv preprint arXiv:2012.11951*, 2020.
- Zhiheng Xi, Wenxiang Chen, Xin Guo, Wei He, Yiwen Ding, Boyang Hong, Ming Zhang, Junzhe Wang, Senjie Jin, Enyu Zhou, et al. The rise and potential of large language model based agents: A survey. *arXiv preprint arXiv:2309.07864*, 2023.
- Shunyu Yao, Howard Chen, John Yang, and Karthik Narasimhan. Webshop: Towards scalable real-world web interaction with grounded language agents. *Advances in Neural Information Processing Systems*, 35:20744–20757, 2022a.
- Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. React: Synergizing reasoning and acting in language models. *ArXiv preprint*, abs/2210.03629, 2022b.
- Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. *arXiv preprint arXiv:2305.10601*, 2023.
- Yining Ye, Xin Cong, Yujia Qin, Yankai Lin, Zhiyuan Liu, and Maosong Sun. Large language model as autonomous decision maker. *arXiv preprint arXiv:2308.12519*, 2023.
- Zapier. Zapier — automation makes you move forward. URL <https://zapier.com/>.
- Yue Zhang, Yafu Li, Leyang Cui, Deng Cai, Lemao Liu, Tingchen Fu, Xinting Huang, Enbo Zhao, Yu Zhang, Yulong Chen, et al. Siren’s song in the ai ocean: A survey on hallucination in large language models. *arXiv preprint arXiv:2309.01219*, 2023.