LLMs for Intelligent Automation - Insights from a Systematic Literature Review

Research Paper

David Sonnabend[1], Mahei Manhai Li[1,2] and Christoph Peters[3]

[1] Universität Kassel, ITeG, Kassel, Germany
{david.sonnabend,mahei.li}@uni-kassel.de
[2] Universität St. Gallen, IWI, St. Gallen, Switzerland
{mahei.li}@unisg.ch
[3] Universität der Bundeswehr München, München, Germany
{christoph.peters}@unibw.de

Abstract: Intelligent Automation (IA) aims to overcome the limitations of traditional Robotic Process Automation (RPA) by integrating Artificial Intelligence (AI). One AI technology that promises to address many RPA limitations, such as dealing with unstructured data or changes in the workflow, is Large Language Models (LLMs). However, how LLMs can advance IA has not been systematically investigated so far. To address this gap, we conduct a systematic literature review to examine how LLMs can advance IA. Our findings reveal that LLMs are primarily used to process complex inputs, generate automation workflows from natural language, and guide goal-oriented GUI navigation. Furthermore, we identify a crucial research gap in combining these different intelligence features and enabling continuous learning at runtime. Thus, we contribute by highlighting opportunities for how LLMs could drive even greater advancements in IA.

Keywords: *Large Language Models (LLMs), Intelligent Process Automation (IPA), Intelligent Automation (IA), Cognitive Automation (CA), Tool Learning*

1 Introduction

To date the commercial value of Large Language Models (LLMs) remains a subject of ongoing debate. While proponents argue that these models have the potential to transform business operations and unlock new efficiencies, sceptics question whether the substantial investment required to develop and deploy them is justified (Dong & Xie, 2024; van Dijk et al., 2023). One area where LLMs could serve as a significant value creator is business process automation (BPA), as their ability to automate tasks traditionally performed by humans can enhance operational efficiency and drive cost reductions (Wornow et al., 2024). If BPA is achieved through any form of 'Artificial Intelligence' (AI), it is often referred to as Intelligent Automation (IA). Accordingly, this paper seeks to answer the following research question: *How can LLMs advance IA?*

To date, BPA often relies on Robotic Process Automation (RPA), which enables the automation of rule-based, repetitive tasks, by mimicking human interactions with software (Chugh et al., 2022; Syed et al., 2020). RPA platforms like Automation Anywhere accomplish this by using visual interfaces and pre-built components to design workflows that execute tasks without traditional coding, and by leveraging recorders to capture user actions. (Chakraborti et al., 2020; Chugh et al., 2022; P et al., 2024). However, traditional RPA has notable limitations, particularly in handling unstructured data, adapting to dynamic environments, and performing complex reasoning tasks (Chen et al., 2009; Ng et al., 2021; Wornow et al., 2024). Furthermore, while RPA promises a low-code approach, in practice, it often relies on highly skilled workers to design, implement and monitor the RPA workflows effectively (Chakraborti et al., 2020; Wornow et al., 2024).

To address some of these shortcomings, the concept of IA was developed, encompassing all efforts to integrate AI into BPA, specifically enhancing RPA with AI capabilities (Chakraborti et al., 2020; Ng et al., 2021; Siderska et al., 2023; Williams & Olajide, 2022). This includes a variety of technologies, ranging from support vector machines to natural language processing techniques (Ng et al., 2021).

However, since the release of ChatGPT in 2022, the emergence of powerful LLMs has transformed our perception of AI. With their advanced natural language understanding, generative capabilities, and contextual reasoning (Minaee et al., 2024), we expect that LLMs will drive a paradigm shift in automation through LLM-powered RPA, representing a new realization of IA. This expectation is further reinforced when considering the extension of LLM capabilities through methods like tool learning (Qin et al., 2025; Qu et al., 2025) and within agent frameworks (Wang et al., 2024).

While there are some literature reviews available on IA, none focus on the specific capabilities that LLMs bring to this domain (Afrin et al., 2025; Patrício et al., 2024; Siderska et al., 2023). Against this backdrop, with this Systematic Literature Review, we aim to analyze how LLMs can advance IA. This is relevant for both theory and practice, as practitioners gain an overview of existing opportunities and challenges in LLM-driven IA, while scientists will be presented with a conceptual framework for understanding LLMs in IA, an overview on current research trajectories and some potentially fruitful future research avenues (Kitchenham, 2004).

Our findings reveal that in IA, LLMs are primarily used for processing complex inputs, generating automation workflows from natural language, and facilitating goal-oriented GUI navigation. However, the integration of these diverse intelligent features, as well as enabling continuous learning at runtime, remains a research gap.

2 Related Work

In this section, we aim to briefly discuss the concept of IA and related notions. Given that our literature review emphasizes the role of LLMs in IA, we will also clarify their primary technological capabilities.

2.1 The Concept of Intelligent Automation

While we use the term IA in this paper, it is important to note that two closely related concepts exist: First, the distinction between IA and Intelligent Process Automation (IPA) should be noted. Researchers like Ng et al. (2021) view IA as the broader concept, with IPA representing a developmental step in the evolution from RPA to autonomous agents, characterized by increased cognitive capabilities and process complexity (Ng et al., 2021). Another term often used for referring to IA is Cognitive Automation (CA). Although some authors emphasize the distinction between cognition and intelligence (Engel et al., 2022), we believe that, for practical purposes, the terminologies of CA and IA can be used interchangeably. Regarding terminology, we further acknowledge the distinctions between the term's *tasks*, *workflows*, and *processes*, which denote increasingly broader scopes of action sequences (Georgakopoulos et al., 1995; Klessascheck et al., 2024); however, throughout this paper, we try to adhere to the terminology used by the authors of the works discussed.

IA itself is frequently defined simply as the integration of AI into RPA (Chakraborti et al., 2020; Ng et al., 2021; Siderska et al., 2023; Williams & Olajide, 2022). However defining AI is challenging, mirroring the unresolved definition of intelligence in psychology (Legg & Hutter, 2007a, 2007b; Wang, 2019). While some researchers define IA narrowly as technology replacing human learning and problem-solving (Coombs et al., 2020), this overlooks core components of intelligence: adaptation to environmental changes (e.g., Sternberg, 1997) and information processing (Hunt, 1980). To address this, we propose a working definition of IA centered on four key capabilities identified in the literature that an IA system should possess: processing complex inputs, adapting to environmental changes, learning from feedback, and logical reasoning. Later, we will systematize how LLMs can advance IA across these capabilities.

2.2 Capabilities of Large Language Models

Following Minaee et al., (2024) LLMs demonstrate capabilities across three functional tiers. Those are, first, basic functionalities, which include world knowledge, text comprehension, multilingual fluency, and code generation; second, emerging capabilities, encompassing instruction following, in-context learning, and reasoning, allowing models to generalize across various tasks; and finally augmented functionalities which leverage external tools, enhance user interaction, and support self-improvement mechanisms (Minaee et al., 2024). Furthermore, LLMs are increasingly discussed to adopt agentic roles, which includes autonomously using tools and replicating human-level decision processes (L. Wang et al., 2024). Among other this trend is driven by Tool Learning paradigms, aiming for LLMs to dynamically acquire, select, and apply external tools to solve tasks more efficiently (Qin et al., 2025; Qu et al., 2025).

3 Method

To systematically explore the existing literature on how LLMs can enhance IA, we conducted a Systematic Literature Review following the PRISMA guidelines (Page, McKenzie, et al., 2021; Page, Moher, et al., 2021). Our search strategy was designed to capture a broad spectrum of relevant studies by using the following search terms: *"Process Automation" AND ("Tool Learning" OR "Function Calling")*, *(Business Process Automation OR Robotic Process Automation) AND (Foundation Models OR Large Language Models OR Agents OR Artificial Intelligence OR Agentic AI OR Gen.AI OR Generative AI)*, as well as *"Cognitive Automation"*, *"Smart Robotic Process Automation"*, *"Agentic Process Automation"*, *"Intelligent Automation"*, and *"Intelligent Process Automation"*.

We conducted a comprehensive search for English-language, peer-reviewed papers from journals and conferences across six academic databases: AISel, ScienceDirect, IEEE Xplore, EBSCO Business, ACM Digital Library, and ProQuest. This search was conducted between February 20, 2025, and March 14, 2025. Given that the release of ChatGPT in 2022 marked a significant turning point for the practical application of LLMs in automation, we restricted our search to studies published from 2022 onward. The initial search across the selected databases yielded 2,085 hits. After excluding records irrelevant to our research question on how LLMs can advance IA, based on title screening and, in cases of doubt, short abstract screening, we narrowed this down to 109 reports. However, 31 of these reports were not retrieved due to redundancy between different search strings. Of the remaining 78 reports, we excluded a further 62 for not aligning with our research question, because they either referred to techniques like RPA as AI without considering their integration with LLMs or merely suggested that LLMs could impact automation without explaining how.

To complement our initial database search, we also conducted a backward and forward search, where we considered two papers from the backward search but then discarding them for having no direct contribution to the process automation field. Additionally, using the same search strings as with the databases, we conducted a search on Google Scholar, which resulted in the inclusion of three additional preprints.

Finally, we read each study in detail and categorized them based on their contributions to the previously identified AI capabilities: the ability to process complex inputs, adapt to environmental changes, learn from feedback, and logically solve problems (see Table 1). Additionally, we remained open to the inductive emergence of new themes to thoroughly address our qualitative 'how' research question.

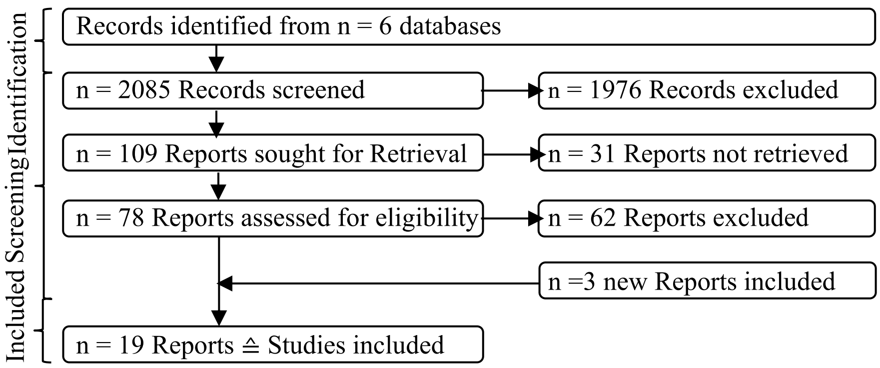

Figure 1. PRISMA diagram adapted from Page et al. (2021)

4 Findings

First, we found that all prototypes and empirical studies in this review fell into at least one of three categories based on their functional use of LLMs. These were: I) LLMs as Tools: The LLM was used for a specific task within the automation workflow. II) LLMs to Build IA: LLMs were used to build or improve automation workflows. III) LLMs as Agents: LLMs were used to steer the IA workflow during execution, such as deciding which tools to use. Finally, it is worth noting that most of the studies did not refer to themselves as IA, IPA or CA, but rather discussed combining LLMs with RPA.

Table 1. Concept Matrix

	Haase et al. (2024)	Jasińska, Lewicz, & Rostalski (2023)	Lo et al. (2024)	Abdellaif, Hassan, & Hamdi (2024)	Gal-Nadasan et al. (2024)	Lin (2024)	Latif et al. (2024)	Lin & Zheng (2024)	Sufi (2025)	Dan et al. (2022)	Bavaresco et al. (2023)	Bandlamudi et al. (2024)	Duesterwald et al. (2024)	Nakagawa, Nitta, & Tsuchiya (2024)	Zeng et al. (2023)	Datta et al. (2024)	Jain et al. (2024)	Wornow et al. (2024)	Ye et al. (2023)
Method																			
Prototype				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Empirical		X																	
Conceptual	X	X	X																
IA Capabilities																			
Complex Input		X	X	X	X	X	X	X	X	X	X								X
Adoption																X	X	X	X
Learning																			
Reasoning			X				X												X
Role of LLMs																			
as Tool		X	X	X	X	X	X	X	X	X	X								X
to build IA											X	X	X	X	X				X
as Agents													X			X	X	X	X

4.1 LLMs as Tools

In this section, we will present several studies that leverage LLMs to solve specific tasks within an automation workflow. These tasks include among others document understanding, search, digitalizing handwritten documents and providing a chat-based interface for RPA solutions, showcasing the versatile applications of LLMs within automation workflows. Therefore, most of the papers presented in this section will mostly contribute to achieving the IA capability to process complex input.

First, Jasińska, Lewicz, and Rostalski (2023) combined a literature review with a case study to identify use cases within a company, including text translation for global communication, enhancing HR processes like recruitment, analyzing vast amounts of web text data, streamlining email management through mailbox automation, monitoring social media for real-time responsiveness, and supporting daily work tasks to improve efficiency and reduce human error (Jasińska, Lewicz, and Rostalski 2023). Furthermore, Lo et al. (2024) also demonstrated in three case studies how LLMs can support automation by analyzing unstructured data (e.g., medical records, customer feedback) and generating actionable recommendations (e.g., treatment plans, demand forecasts) (Lo et al. 2024). Thus, their work not only explores the capability of processing complex input but also the ability to reason about a problem.

That those endeavors are also technically achievable was shown multiple times: First, it was shown that it is possible to build workflows that leverage RPA for document handling (e.g., gathering, opening, uploading) and LLMs for generating accurate text summaries (Lin, 2024). In another study, it was demonstrated that integrating RPA with LLMs can effectively grade student assignments within a digital learning platform, again using RPA technology for navigating the digital learning platform and the LLM to grade the student assignments given a prompt specifying how (Latif et al., 2024). Hence, this study successfully utilized LLMs not only for handling complex input (text) but also for reasoning about a problem.

Another identified use case of the IA capability to process complex data is the digitization of handwritten documents. It was demonstrated that the accuracy and efficiency of data extraction from handwritten and unstructured document scans could be enhanced by not only utilizing state-of-the-art optical character recognition (OCR) but also by forwarding the extracted information to an LLM. The LLM was tasked with interpreting the extracted data and storing it in JSON format. The developed prototype was validated in the context of immigration services and showed superior performance compared to existing RPA applications (Abdellaif et al., 2024). Conceptually, this approach closely resembles the demonstrated use of OpenAI's GPT-4 to extract data from handwritten medical forms within an RPA workflow, however without relying on an intermediary third-party OCR module (Gal-Nadasan et al., 2024).

Two other studies at the intersection of complex data processing and reasoning combined RPA techniques with LLMs for automating search tasks. To this end, Sufi (2025) developed a chatbot for real-time personalized news delivery. Within an RPA workflow, a LLM was used to generate queries against a news database, demonstrating the ability to accurately interpret user intent and translate it into search queries (Sufi, 2025). A similar configuration also proved effective in executing a broad range of web search tasks (Lin & Zheng, 2024).

Another emerging trend we observed in the use of LLMs within BPA is the concept of addressing RPA workflows through chats. While it might be debatable whether this approach fully aligns with our categorization of LLMs as tools within RPA workflows, we will briefly highlight two studies that employ this design here: The first study, conducted in the eCommerce sector, leveraged the DRUID Chatbot Platform and UiPath RPA. This solution enabled consumers to interact with chatbots across multiple channels, including Web, WhatsApp, and Facebook Messenger, for product exploration and purchases, while RPA automates invoice generation and distribution (Dan et al., 2022). The second study demonstrated in a field study with Dell Inc. that a chatbot can translate unstructured user requests into structured RPA inputs. By mimicking natural human communication, the chatbot interface reduced the need for employees to learn technical workflows, thereby lowering barriers to adoption and enhancing overall efficiency (Bavaresco et al., 2023). In these two studies, the capability of LLMs to process complex input, such as natural language, was hence utilized to lower the barrier to interacting with automation solutions, a theme that will recur in the next section.

4.2 LLMs to build IA

Another role that LLMs can play in advancing IA is their ability to build and improve automation workflows themselves. By merging their capacity to process complex inputs with their inherent reasoning skills, LLMs can transform unstructured input into structured, formalized automation workflows. This represents a shift toward the automation of automation, where LLM driven IA systems contribute to the design and optimization of automation processes. However, it is important to clarify that the workflows developed and improved in the following studies remained within the capabilities of conventional RPA, therefore not inherently demonstrating advanced IA capabilities.

Two studies on leveraging LLMs for building RPA workflows were conducted by research teams at IBM Research. First, Bandlamudi et al. (2024) propose a framework, which leverages LLMs during the build phase of chat-based RPA-Solutions to generate conversational artifacts—such as training data for intent classification, slot-filling questions, and next-best action recommendations—directly from API specifications. This automates expertise- and labor-intensive tasks like utterance generation and semantic input-output matching (training data for intend classification), while retaining human oversight to ensure domain-specific accuracy. A key advancement is the framework’s automated testing mechanism, where LLMs act as proxy users and judges to simulate and evaluate conversations. The system’s architecture, validated across datasets and deployed in IBM Watson, is reported to demonstrate tangible improvements: intent classification accuracy gains, context-aware next best agent recommendations outperforming keyword-based methods, and scalable integration with platforms like Salesforce and SAP (Bandlamudi et al., 2024). In contrast, the second IBM study focused on the actual workflow generation rather than the ability of these workflows to be addressed via natural language. Duesterwald et al. (2024) proposed a framework, built into Watsonx Orchestrate, that uses LLMs to convert natural language into executable workflows via a Python-based intermediate representation. They claim this avoids error-prone XML/JSON outputs, leveraging LLMs’ code-generation strengths to reduce syntax errors and streamline automation. They developed a multi-agent system that handles tasks like API mapping and workflow refinement, coordinated by an LLM orchestrator that dynamically routes tasks using in-context learning instead of rigid rules (Duesterwald et al., 2024). Similarly focusing on workflow generation but differing in methodology, JP Morgan AI Research presents FlowMind, a generic prompt recipe for using LLMs to build workflows on APIs (Zeng et al., 2023). While Duesterwald et al. (2024) emphasize structured code conversion within a dedicated automation environment, FlowMind is designed as a more flexible prompting technique, applicable across different API-based workflows. The authors claim their approach mitigates the common issue of hallucinations in LLMs and eliminates direct interaction between LLMs and proprietary data or code. The recipe includes setting up the context, enumerating available APIs with function declarations, parameters, and descriptions, and prompting the LLM to write workflow code using these APIs. The LLM then generates workflow code using the introduced APIs, which is executed to provide output to the user. Additionally, a feedback mechanism allows the system to present a high-level description of the generated workflow to the user, who can then provide feedback to refine it (Zeng et al., 2023).

While the previous studies focus on workflow generation and automation efficiency, Nakagawa, Nitta, and Tsuchiya (2024) take a distinct approach by addressing error detection in RPA workflows. To this end they first systematically categorize common RPA issues, such as substitutable processes, access violations, and inappropriate argument values, based on an analysis of 87 bots developed using *Automation Anywhere*. To automate defect detection, they leverage LLMs, specifically Word2Vec and Word Mover's Distance, to calculate the similarity between workflow code and previously identified issue classes, as well as to extract key features indicative of bot smells. They claim, their tool effectively identifies defects with high accuracy, addressing challenges unique to RPA, such as UI-based interactions and low-code development constraints, and therefore significantly enhances RPA maintainability and debugging efficiency (Nakagawa et al., 2024).

4.3 LLMs as Agents

Finally**,** we identified four papers that utilize LLMs with RPA in an agentic manner, enabling LLMs to make decisions and use tools within the IA workflow. Notably, these architectures frequently depend on constructing workflows at runtime, linking these studies to the research presented in 'LLMs to Build IA,' while also proposing a new paradigm. A key distinction among them lies in their method of interaction with other software: three studies emphasize GUI-based task execution (Datta et al., 2024; Jain et al., 2024; Wornow et al., 2024), whereas one adopts a code-based orchestration approach (Ye et al., 2023).

Datta et al. (2024) propose AUTONODE a novel methodology that integrates LLMs with computer vision to enhance RPA. Initially, the system captures screenshots of the GUI and processes them using advanced object detection and optical character recognition to extract relevant elements. These visual inputs are then fed into an LLM, which functions as an autonomous agent by interpreting the current state, deciding the next optimal action (such as clicking, typing, or scrolling), and executing it. To mitigate issues like imprecise targeting and action errors, the framework evolved through iterative refinements: the first iteration relied solely on visual processing and LLM-based decision-making, the second introduced explicit contextual instructions and a verification step to reduce errors, and the third adopted a graph-based approach. In this final iteration, the system organizes interface elements into a structured knowledge graph that mirrors human attention by prioritizing regions of interest, enabling the LLM to navigate and make decisions more accurately through graph traversal and similarity assessments. This final configuration achieved near-human success rates, outperforming existing tools like MultiOn. The authors claim this hybrid approach—melding LLMs, computer vision, and graph-based logic—represents a paradigm shift toward self-learnable, *cognitive RPA* that dynamically adapts to real-world applications like email management or calendar scheduling, bridging the gap between rule-based automation and human-like problem-solving (Datta et al., 2024).

ECLAIR (Wornow et al., 2024) follows a comparable GUI-driven strategy but differentiates itself through its focus on learning from demonstrations and self-validation. First, Vision Language Models document workflows by analyzing visual demonstrations (key video frames) and textual documentation, achieving high accuracy rates in generating step-by-step procedures through screenshot-based reasoning and alignment with action logs. Second, during execution, LLMs decompose workflows into GUI actions: This includes action planning using documentation-guided reasoning to predict next steps (e.g., "click submit"), and mapping these steps to GUI elements (e.g., HTML element labelling). Additionally, Wornow et al. (2024) propose a self-validation procedure that uses LLMs to monitor workflow execution at two levels: (1) the model checks if individual actions (e.g., clicking a button) are feasible and successful under certain constraints (e.g., "Is the button visible?"), and (2) the model assesses whether the overall workflow was completed successfully and aligns with the intended trajectory. Though evaluation on a subset of 30 workflows from the WebArena benchmark (Zhou et al., 2024) showed that ECLAIRE still suffered from some shortcomings regarding low-level execution precision, the authors claim that ECLAIR is a paradigm shift from static RPA to AI-driven, context-aware automation, through bypassing RPA's rigid rule-coding, adapting to GUI variations, and minimizes setup/maintenance costs (Wornow et al., 2024).

The last paper focusing on workflow execution through LLM-driven GUI navigation introduces a system called SmartFlow. Like earlier research, it combines LLMs to interpret HTML code with vision models for layout mapping. SmartFlow generates action sequences from the extracted information, executing them via a scripting engine to complete tasks. The authors assert that this approach enhances adaptability to GUI changes and effectively manages complex tasks. Although SmartFlow demonstrated strong performance on a dataset of varied enterprise applications, developed by the authors, it faces limitations, such as difficulties in handling dynamic fields (Jain et al., 2024).

Taken together, these three papers show different implementations of how language models can be utilized to navigate GUI interactions in automation workflows. By actively interpreting the GUI at runtime instead of relying on predefined click positions, they fulfill the IA characteristic of adaptability, addressing the core RPA limitation of being too sensitive to GUI changes.

Ye et al. (2023) take a fundamentally different approach with ProAgent, shifting away from GUI-based navigation toward workflow construction through code. ProAgent employs a proposed agentic workflow description language, which uses JSON for standardizing data flow and Python for defining control flow. Its architecture incorporates two specialized agents, DataAgent and ControlAgent: DataAgent is designed to handle complex data processing tasks, autonomously executing actions based on natural language task descriptions. ControlAgent, on the other hand, functions as a dynamic controller, making real-time decisions based on input data to determine the subsequent actions in the workflow. The workflow construction process in ProAgent is iterative, involving operations such as action definition, implementation, and workflow orchestration, culminating in a task submission to finalize the workflow. During execution, ProAgent employs a Python interpreter to sequentially execute the workflow. Thus, ProAgent integrates the use of LLMs as agents, tools, and for constructing workflows, distinguishing it from other studies presented in this section (Ye et al., 2023).

5 Discussion

First, we must state that the amount of research at the intersection of LLMs and process automation remains relatively scarce. A possible explanation for this gap could be, despite the long-established concept of IA, the fundamental tension between the traditionally deterministic nature of RPA and the inherently probabilistic behavior of LLMs (Siderska et al., 2023).

Despite this scarcity, we believe this review offers valuable insights for both academic and practice, which we will discuss in the following paragraphs. Afterward, we will highlight and explore some identified research gaps in depth.

5.1 Implications

Implications for Practice. The integration of LLMs as tools into RPA workflows to advance IA offers businesses numerous automation opportunities and is technically straightforward, as demonstrated by many successful implementations. (e.g., Abdellaif et al., 2024; Jasińska et al., 2023; Lin and Zheng, 2024). A particularly interesting use case in this area could be the communication with RPA workflows through chatbots (Bavaresco et al., 2023; Dan et al., 2022). Another area where LLMs promises to have a profound impact for IA is the substantial reduction of some major downsides usually associated with RPA. First, the creation and maintenance of RPA workflows, which traditionally require substantial manual effort, can now be significantly simplified by building workflows directly from natural language (Bandlamudi et al., 2024; Duesterwald et al., 2024; Zeng et al., 2023). A second major improvement that LLMs can provide in IA is the increased flexibility and robustness. Traditional RPA workflows tend to be rigid and highly sensitive to changes in input data, workflow, or GUI. LLMs can process complex input data and reason about the next steps necessary to achieve a goal (Ye et al., 2023). Furthermore, LLMs combined with computer vision techniques can mitigate this issue by interpreting modifications in the GUI and adapting navigation strategies goal oriented accordingly (Datta et al., 2024; Jain et al., 2024; Wornow et al., 2024). Despite these advancements, challenges in GUI navigation persist (Jain et al., 2024; Wornow et al., 2024), underscoring the need for software providers to make their applications more accessible through API interfaces to ensure their systems are both human- and AI-friendly.

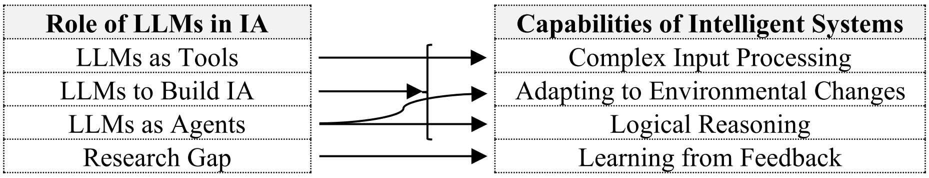

Figure 2. Building IA with LLMs

Implications for Theory. Within our literature review, we categorized how LLMs can be utilized in IA, identifying three primary roles: LLMs as tools, LLMs for building IA, and LLMs as agents within automation. This categorization may assist researchers in contextualizing their work. Furthermore, there seems to be a mapping between these roles and the intelligent features realized in the respective implementations (See Figure 2): LLMs as tools are predominantly employed to process complex inputs, such as unstructured text documents. LLMs as agents, on the other hand, are utilized to enhance adaptability and reason within workflows. Additionally, we observe that systems combining the power of LLMs with more formal structures, such as graphs, are often the most effective (Datta et al., 2024). Furthermore, the integration of LLMs in IA has yet to realize the core capability of intelligence to learn from feedback. This shortfall likely arises from the common practice of deploying LLMs as static models, since real-time fine-tuning is difficult to implement (Minaee et al., 2024). However, it is important to note that mechanisms like information retrieval could enable dynamic learning by supplying relevant context in real time without requiring retraining (Minaee et al., 2024). Finally, when LLMs are used to build automation workflows, they can in principle support all intelligent capabilities of IA, but since this occurs outside of runtime, their role is structurally distinct from LLMs as tools or agents.

5.2 Future Research

While numerous studies explore the management implications of IA in general, none explicitly focus on the unique challenges and opportunities introduced by LLMs for IA. First, as LLMs significantly lower the barriers to building automation workflows, future research should examine how this affects the adoption dynamics of IA (Dirnberger-Wild & Roth, 2024; Engel et al., 2022; Mayr et al., 2024). Second, LLMs' ability to process more complex data structures and execute advanced decision-making within IA workflows could intensify known risks associated with IA, such as skill erosion (Rinta-Kahila et al., 2023), security vulnerabilities (Al-Slais & Ali, 2023), and in business continuity (Brás et al., 2023). While some of these emerging research questions have been theoretically discussed and conceptualized in the context of LLMs (Haase et al., 2024), no empiric answers have been presented so far.

Future research should also systematically evaluate the usability of smaller LLMs in IA systems, considering their potential environmental (Bossert & Loh, 2025) and financial benefits. While large-scale models offer powerful capabilities, smaller models could provide more sustainable and cost-effective alternatives, making IA adoption more accessible across industries (Irugalbandara et al., 2024). Additionally, it may be worthwhile to investigate how human-automation interaction (Vu et al., 2023) evolves as LLMs enable IA to become more 'intelligent'.

Finally, our review highlights a surprising lack of studies that attempt to combine multiple IA capabilities. Integrating these different IA capabilities could enable systems to tackle even more complex tasks, allowing for goal-driven automation of longer processes. Additionally, we identified a gap in proposed LLM-driven IA architectures that actively learn from feedback—an essential dimension of intelligence. To address this, we propose a deeper exploration into a novel AI paradigm: Tool Learning.

Tool learning enables LLMs to dynamically acquire, select, and apply external tools for more efficient task solving (Qin et al., 2025; Qu et al., 2025). A key aspect of tool learning is its ability to integrate multiple tools in parallel or sequentially, optimizing for efficiency and accuracy. Beyond selecting and using existing tools, these systems may also engage in tool creation, where new tools are dynamically constructed to meet specific problem-solving needs (Qin et al., 2025). This aligns with developments in LLM-based agents, which exhibit autonomy in refining strategies and integrating novel capabilities into their workflows (Wang et al., 2024).

Effectively utilizing this paradigm could significantly advance IA towards a scenario where automation involves continuous learning and process optimization. Existing LLM-driven GUI navigation approaches could be encapsulated as independent tools, allowing automation systems to engage in tool creation as part of their operational framework. By enabling LLMs to generate and refine tools dynamically, these systems could evolve beyond workflow execution into more autonomous and adaptive automation paradigms. We envision a system that actively researches how to solve higher-level tasks by searching internal documentation, software documentation, and the internet, thereby representing the long-anticipated shift from automating processes to automating jobs and proving the true commercial value of LLMs.

5.3 Conclusion

In this study, we presented numerous studies that leverage LLMs to process complex inputs in IA. Additionally, our findings highlight an emerging trend of using natural language to address and build automated workflows, enhancing both adaptability and usability. Furthermore, we presented research that utilizes LLMs to orchestrate workflows and improve adaptability by enabling goal-oriented GUI navigation.

Despite these advancements, our review identified a significant research gap in integrating all four key features of IA within one IA system: complex input processing, environmental adaptability, learning from feedback, and reasoning. Notably, we observed a complete absence of mechanisms for learning from feedback, which remains a critical limitation in current approaches.

References

Abdellaif, O. H., Hassan, A. N., & Hamdi, A. (2024). ERPA: Efficient RPA Model Integrating OCR and LLMs for Intelligent Document Processing. *2024 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*, 295–300. https://doi.org/10.1109/MIUCC62295.2024.10783599

Afrin, S., Roksana, S., & Akram, R. (2025). AI-Enhanced Robotic Process Automation: A Review of Intelligent Automation Innovations. *IEEE Access*, **13**, 173–197. https://doi.org/10.1109/ACCESS.2024.3513279

Al-Slais, Y., & Ali, M. (2023). Robotic Process Automation and Intelligent Automation Security Challenges: A Review. *2023 International Conference On Cyber Management And Engineering (CyMaEn)*, 71–77. https://doi.org/10.1109/CyMaEn57228.2023.10050996

Bandlamudi, J., Mukherjee, K., Agarwal, P., Chaudhuri, R., Pimplikar, R., Dechu, S., Straley, A., Ponniah, A., & Sindhgatta, R. (2024). Framework to enable and test conversational assistant for APIs and RPAs. *AI Magazine*, **45**(4), 443–456. https://doi.org/10.1002/aaai.12198

Bavaresco, R. S., Nesi, L. C., Victória Barbosa, J. L., Antunes, R. S., Da Rosa Righi, R., Da Costa, C. A., Vanzin, M., Dornelles, D., Junior, S. C., Gatti, C., Ferreira, M., Silva, E., & Moreira, C. (2023). Machine learning-based automation of accounting services: An exploratory case study. *International Journal of Accounting Information Systems*, **49**, 100618. https://doi.org/10.1016/j.accinf.2023.100618

Bossert, L. N., & Loh, W. (2025). Why the carbon footprint of generative large language models alone will not help us assess their sustainability. *Nature Machine Intelligence*, *7*(2), 164–165. https://doi.org/10.1038/s42256-025-00979-y

Brás, J., Pereira, R., & Moro, S. (2023). Intelligent Process Automation and Business Continuity: Areas for Future Research. *Information*, *14*(2), 122. https://doi.org/10.3390/info14020122

Chakraborti, T., Isahagian, V., Khalaf, R., Khazaeni, Y., Muthusamy, V., Rizk, Y., & Unuvar, M. (2020). *From Robotic Process Automation to Intelligent Process Automation: Emerging Trends* (No. arXiv:2007.13257). arXiv. https://doi.org/10.48550/arXiv.2007.13257

Chen, X., Liu, W., & Williams, M.-A. (2009). Introduction: Practical Cognitive Agents and Robots. *Autonomous Agents and Multi-Agent Systems*, *19*(3), 245–247. https://doi.org/10.1007/s10458-009-9077-x

Chugh, R., Macht, S., & Hossain, R. (2022). Robotic Process Automation: A review of organizational grey literature. *International Journal of Information Systems and Project Management*, **10**(1), 5–26. https://doi.org/10.12821/ijispm100101

Dan, G., Claudiu, D., Alexandra, F., Adela, P.-P., Ovidiu, S., Honoriu, V., & Liviu, M. (2022). Multi-Channel Chatbot and Robotic Process Automation. *2022 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*, 1–6. https://doi.org/10.1109/AQTR55203.2022.9801960

Datta, A., Verma, T., Chawla, R., S, M. N., & Bhola, I. (2024). AUTONODE: A Neuro-Graphic Self-Learnable Engine for Cognitive GUI Automation. *2024 29th International Conference on Automation and Computing (ICAC)*, 1–7. https://doi.org/10.1109/ICAC61394.2024.10718742

Dirnberger-Wild, J., & Roth, M. (2024). Conceptual Framework Introducing the Success Factors for Implementing Intelligent Automation-A Qualitative Multiple Case Study. *Proceedings of the 2024 10th International Conference on Computer Technology Applications*, 8–14. https://doi.org/10.1145/3674558.3674560

Dong, H., & Xie, S. (2024). *Large Language Models (LLMs): Deployment, Tokenomics and Sustainability* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2405.17147

Duesterwald, E., Isahagian, V., Jayaram, K. R., Kumar, R., Muthusamy, V., Oum, P., Thomas, G., & Venkateswaran, P. (2024). A Conversational Assistant Framework for Automation. *Proceedings of the 25th International Middleware Conference Industrial Track on ZZZ*, 1–7. https://doi.org/10.1145/3700824.3701093

Engel, C., Ebel, P., & Leimeister, J. M. (2022). Cognitive automation. *Electronic Markets*, **32**(1), 339–350. https://doi.org/10.1007/s12525-021-00519-7

Gal-Nadasan, N., Stoicu-Tivadar, V., Gal-Nadasan, E., & Raluca Dinu, A. (2024). Handwritten Data Extraction Using OpenAI ChatGPT4o and Robotic Process Automation. In L. Stoicu-Tivadar, A. Benis, T. M. Deserno, S. D. Bolboacă, K. Saranto, M. Crişan-Vida, P. Gallos, O. S. Chirila, P. Weber, G. Mihalaş, & O. Tamburis (Eds.), *Studies in Health Technology and Informatics*. IOS Press. https://doi.org/10.3233/SHTI241101

Georgakopoulos, D., Hornick, M., & Sheth, A. (1995). An overview of workflow management: From process modeling to workflow automation infrastructure. *Distributed and Parallel Databases*, **3**(2), 119–153. https://doi.org/10.1007/BF01277643

Haase, J., Kremser, W., Leopold, H., Mendling, J., Onnasch, L., & Plattfaut, R. (2024). Interdisciplinary directions for researching the effects of robotic process automation and large language models on business processes. *Communications of the Association for Information Systems*, **54**(1), 28. https://doi.org/10.17705/1CAIS.05421

Hunt, E. (1980). Intelligence as an information-processing concept. *British Journal of Psychology*, *71*(4), 449–474. https://doi.org/10.1111/j.2044-8295.1980.tb01760.x

Irugalbandara, C., Mahendra, A., Daynauth, R., Arachchige, T. K., Dantanarayana, J., Flautner, K., Tang, L., Kang, Y., & Mars, J. (2024). Scaling Down to Scale Up: A Cost-Benefit Analysis of Replacing OpenAI's LLM with Open Source SLMs in Production. *2024 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, 280–291. https://doi.org/10.1109/ISPASS61541.2024.00034

Jain, A., Paliwal, S., Sharma, M., Vig, L., & Shroff, G. (2024). *SmartFlow: Robotic Process Automation using LLMs* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2405.12842

Jasińska, K., Lewicz, M., & Rostalski, M. (2023). Digitization of the enterprise—Prospects for process automation with using RPA and GPT integration. *Procedia Computer Science*, **225**, 3243–3254. https://doi.org/10.1016/j.procs.2023.10.318

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, **33**(2004), 1-26.

Klessascheck, F., Bein, L., Haase, J., & Pufahl, L. (2024). A Critical Investigation of Rationalities in Automation with BPM. *2024 26th International Conference on Business Informatics (CBI)*, 30–39. https://doi.org/10.1109/CBI62504.2024.00014

Latif, A., ElBarachi, M., Tsaramirsis, G., & Khan, F. Q. (2024). Generative AI Enabled Robotic Process Automation: A Practical Case Study. *2024 IEEE 30th International Conference on Parallel and Distributed Systems (ICPADS)*, 769–775. https://doi.org/10.1109/ICPADS63350.2024.00104

Legg, S., & Hutter, M. (2007a). *A Collection of Definitions of Intelligence* (No. arXiv:0706.3639). arXiv. https://doi.org/10.48550/arXiv.0706.3639

Legg, S., & Hutter, M. (2007b). Universal Intelligence: A Definition of Machine Intelligence. *Minds and Machines*, **17**(4), 391–444. https://doi.org/10.1007/s11023-007-9079-x

Lin, M. (2024). Design and implementation of batch file intelligent summary assistant based on AIGC+RPA. *Proceedings of the 2024 8th International Conference on Electronic Information Technology and Computer Engineering*, 551–556. https://doi.org/10.1145/3711129.3711224

Lin, M., & Zheng, Z. (2024). Design and implementation of intelligent search assistant based on AIGC and RPA. *Proceedings of the 2024 8th International Conference on Electronic Information Technology and Computer Engineering*, 534–538. https://doi.org/10.1145/3711129.3711221

Mayr, A., Stahmann, P., Nebel, M., & Janiesch, C. (2024). Still doing it yourself? Investigating determinants for the adoption of intelligent process automation. *Electronic Markets*, **34**(1), 56. https://doi.org/10.1007/s12525-024-00737-9

Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2024). *Large Language Models: A Survey* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2402.06196

Nakagawa, H., Nitta, S., & Tsuchiya, T. (2024). Review-Based Bot Smell Classification in Robotic Process Automation. *Procedia Computer Science*, **246**, 381–390. https://doi.org/10.1016/j.procs.2024.09.417

Ng, K. K. H., Chen, C.-H., Lee, C. K. M., Jiao, J. (Roger), & Yang, Z.-X. (2021). A systematic literature review on intelligent automation: Aligning concepts from theory, practice, and future perspectives. *Advanced Engineering Informatics*, **47**, 101246. https://doi.org/10.1016/j.aei.2021.101246

P, J. A., Shankar, A., Jr, A. N., Maniyar, H., Sr, N., & Ambarkhane, S. (2024). Robotic Process Automation: In-Depth Analysis of Advanced Automation Techniques and Technologies. *2024 IEEE International Conference on Contemporary Computing and Communications (InC4)*, 1–6. https://doi.org/10.1109/InC460750.2024.10649336

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, n160. https://doi.org/10.1136/bmj.n160

Patrício, L., Varela, L., & Silveira, Z. (2024). Integration of Artificial Intelligence and Robotic Process Automation: Literature Review and Proposal for a Sustainable Model. *Applied Sciences*, **14**(21), 9648. https://doi.org/10.3390/app14219648

Plattfaut, R., Borghoff, V., Godefroid, M., Koch, J., Trampler, M., & Coners, A. (2022). The Critical Success Factors for Robotic Process Automation. *Computers in Industry*, *138*, 103646. https://doi.org/10.1016/j.compind.2022.103646

Qin, Y., Hu, S., Lin, Y., Chen, W., Ding, N., Cui, G., Zeng, Z., Zhou, X., Huang, Y., Xiao, C., Han, C., Fung, Y. R., Su, Y., Wang, H., Qian, C., Tian, R., Zhu, K., Liang, S., Shen, X., … Sun, M. (2025). Tool Learning with Foundation Models. *ACM Computing Surveys*, **57**(4), 1–40. https://doi.org/10.1145/3704435

Qu, C., Dai, S., Wei, X., Cai, H., Wang, S., Yin, D., Xu, J., & Wen, J. (2025). Tool learning with large language models: A survey. *Frontiers of Computer Science*, **19**(8), 198343. https://doi.org/10.1007/s11704-024-40678-2

Rinta-Kahila, T., The University of Queensland Business School, Penttinen, E., Aalto University School of Business, Salovaara, A., Aalto University / University of Helsinki, Soliman, W., University of Agder, Norway / University of Jyväskylä, Ruissalo, J., & Aalto University School of Business. (2023). The Vicious Circles of Skill Erosion: A Case Study of Cognitive Automation. *Journal of the Association for Information Systems*, **24**(5), 1378–1412. https://doi.org/10.17705/1jais.00829

Siderska, J., Aunimo, L., Süße, T., Stamm, J. V., Kedziora, D., & Aini, S. N. B. M. (2023). Towards Intelligent Automation (IA): Literature Review on the Evolution of Robotic Process Automation (RPA), its Challenges, and Future Trends. *Engineering Management in Production and Services*, **15**(4), 90–103. https://doi.org/10.2478/emj-2023-0030

Sternberg, R. J. (1997). The concept of intelligence and its role in lifelong learning and success. *American Psychologist*, **52**(10), 1030–1037. https://doi.org/10.1037/0003-066X.52.10.1030

Sufi, F. (2025). Just-in-Time News: An AI Chatbot for the Modern Information Age. *AI*, **6**(2), 22. https://doi.org/10.3390/ai6020022

Syed, R., Suriadi, S., Adams, M., Bandara, W., Leemans, S. J. J., Ouyang, C., Ter Hofstede, A. H. M., Van De Weerd, I., Wynn, M. T., & Reijers, H. A. (2020). Robotic Process Automation: Contemporary themes and challenges. *Computers in Industry*, *115*, 103162. https://doi.org/10.1016/j.compind.2019.103162 van Dijk, B. M. A., Kouwenhoven, T., Spruit, M. R., & van Duijn, M. J. (2023). *Large Language Models: The Need for Nuance in Current Debates and a Pragmatic Perspective on Understanding* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2310.19671

Vu, H., Haase, J., Leopold, H., & Mendling, J. (2023). Towards a Theory on Process Automation Effects. In C. Di Francescomarino, A. Burattin, C. Janiesch, & S. Sadiq (Eds.), *Business Process Management Forum* (Vol. 490, pp. 285–301). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-41623-1_17

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, **18**(6), 186345. https://doi.org/10.1007/s11704-024-40231-1

Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, **10**(2), 1–37. https://doi.org/10.2478/jagi-2019-0002

Williams, O. C., & Olajide, F. (2022). Towards the Design of an Intelligent Automation Framework for Business Processes. *2022 5th International Conference on Information and Computer Technologies (ICICT)*, 13–17. https://doi.org/10.1109/ICICT55905.2022.00010

Wornow, M., Narayan, A., Opsahl-Ong, K., McIntyre, Q., Shah, N., & Ré, C. (2024). Automating the Enterprise with Foundation Models. *Proceedings of the VLDB Endowment*, **17**(11), 2805–2812. https://doi.org/10.14778/3681954.3681964

Ye, Y., Cong, X., Tian, S., Cao, J., Wang, H., Qin, Y., Lu, Y., Yu, H., Wang, H., Lin, Y., Liu, Z., & Sun, M. (2023). *ProAgent: From Robotic Process Automation to Agentic Process Automation* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2311.10751

Zeng, Z., Watson, W., Cho, N., Rahimi, S., Reynolds, S., Balch, T., & Veloso, M. (2023). FlowMind: Automatic Workflow Generation with LLMs. *4th ACM International Conference on AI in Finance*, 73–81. https://doi.org/10.1145/3604237.3626908

Zhou, S., Xu, F. F., Zhu, H., Zhou, X., Lo, R., Sridhar, A., Cheng, X., Ou, T., Bisk, Y., Fried, D., Alon, U., & Neubig, G. (2024). *WebArena: A Realistic Web Environment for Building Autonomous Agents* (No. arXiv:2307.13854). arXiv. https://doi.org/10.48550/arXiv.2307.13854