Harvey – An Agentic AI-Based Hr Chatbot Platform for Intelligent Hr Automation

Rucha Salunke, Revati Murgodkar, Nathan Mendis
PES Modern College of Engineering

***Abstract* - HARVEY is an Agentic AI-based HR Chatbot Platform, which automates and boosts the contemporary Human Resource Management (HRM) and is the topic of this paper where the design, development, and evaluation of the platform are discussed. Based on the concepts of agentic artificial intelligence and the generative AI, the system incorporates conversational intelligence with autonomous task completion to facilitate the HR operations. HARVEY combines the reasoning of LLM, Retrieval-Augmented Generation (RAG), and tool-based execution to automate the process of working in HR such as hiring, onboarding, leave management, and employee support. Its design focuses on privacy (supports self-hosted LLMs), explainability (supports action logging and confirmation), and modularity (AgensToolData Provider pattern). Simulated workflow and User Acceptance Testing (UAT) analysis demonstrates significant savings in the process times and high intent-recognition accuracy. The paper summarizes the findings of the extant literature on AI in HR, such as efficiency improvement, redundancy elimination, and the attitude of users, and fills it in with the application of the HARVEY prototype. Using smart natural language interaction, HARVEY does autonomous HR tasks including recruitment support, leave maintenance, onboarding and employee support. The experimental evaluation shows the high efficiency improvement, user experience and reduced query resolution time by 60 percent. The other topics that include implementation issues, ethics, and the future of agentic AI in enterprise human capital management are also explained in this paper.**

***Keywords:* Agentic AI, Human Resources, Chatbot, Automation, NLP, HR Management, AI Integration, Employee Experience, Retrieval-Augmented Generation, Large Language Models**

I. INTRODUCTION

The HR (human resource) departments of industries have been confronted with heightened responsiveness, personalization, as well as operational efficiency demands [1]. The increased number of employees and the mixed work setting has resulted in the escalation of administrative loads, which poses a form of cry of automation and smart support tools [2]. The use of conventional HR activities is still very much reliant on disjointed tools, manual processes, and disintegrated systems that demand repetition of data entry and reiteration of manual data entry and processing [3]. This fragmentation causes inefficiency, slowness in response and low satisfaction among the employees throughout the enterprise [4].

The next step of generative AI is called agentic AI which enables AI systems to perform actions independently in order to accomplish user-specified objectives [5][6]. Contrary to the traditional chatbots, where the system can only respond to the questions, agentic AI systems have the ability to sense their surroundings, think through complex tasks and reason, plan multi-stage tasks, and perform them with minimal guidance [7]. Such systems are based on underlying architectures, like the Transformer [8], or pretrained language representations, like BERT [9], and integrate large language models (LLMs) with planning, memory, and tool execution to allow autonomous orchestration of workflows [10][11].

HARVEY is created to be an intelligent HR assistant that is capable of both conversational AI and autonomous decision-making. The platform curbs the disadvantages of the current HR tools by offering a single interface that can read natural language, storage enterprise databases, and execute HR functions without human intervention [12]. HR departments are bottlenecked in administration and lack strategic capacity; HARVEY applies the principles of agentic AI to an actual HR assistant that SMEs and enterprises will utilize [13]. The idea will be to develop a single HR interface by integrating the LLM-based reasoning with the RAG-based knowledge retrieval and executables in form of modules to develop a full privacy-respecting HR automation platform [14].

This paper has the following contributions: (i) we introduce HARVEY, a new agentic AI-based HR chatbot platform combining LLM reasoning, RAG, and tool-based execution (ii) we present a pattern of an agentic AI-based HR operational efficiency with a modular Agent-Tool-Data Provider architecture (iii) we present empirical findings and user acceptance testing proving the significant increases in HR operational efficiency, and (iv) we comment on ethical issues and future directions of agentic AI used in enterprise human capital management.

II. LITERATURE REVIEW

The integration of Artificial Intelligence into HR management has evolved significantly over the past decade, transitioning from rule-based automation to intelligent, context-aware systems powered by natural language processing and machine learning [15].

2.1 AI in Human Resource Management

Bastida et al. (2025) explored the strategic incorporation of AI in HRM, where the authors note that AI can be used in recruitment, performance assessment, and development and express worries about the issue of algorithmic bias and technostress [1]. Their two-level model is moderate in algorithmic speed and humanitarian in the workforce development. Dima et al. (2024) performed a systematic literature review evaluating the impact of AI on HR activities and the functions of the HR triad (employees, line managers, HR professionals) and found out both opportunities, such as the better decision-making, and challenges, such as perceived fairness issues [16]. Ekuma (2024) conducted an in-depth systematic review of AI and automation in the field of HRD, covering their use in talent development, workforce planning, learning, and performance management [26].

Kshetri (2025) emphasized the use of AI agents to streamline working processes, refine the decision-making process, and transform the experience of employees in the context of redefining HR practices [27]. Kessavane (2025) highlighted the minimization of redundancy and efficiency in HR processes by agentic AI due to intelligent automation [28]. In his master thesis, Suhonen (2025) investigated the attitude and resistance of humans to implementing AI chatbots in HR advisory services, which showed both enthusiasm and opposition of HR professionals towards this solution [29].

2.2 Chatbots and Conversational AI in HR

AI-based chatbots have become area-changing instruments within the HR departments, which can respond to the recruitment questions, onboarding, managing leaves, and answering policy questions [17,18]. Studies show that chatbots save a lot of time and decrease the workload of the administration and increase the satisfaction levels of employees because of 24/7 availability [19]. Nevertheless, the perceived impersonality, as well as the necessity of human escalation pathways in complex situations are also mentioned in the studies [20].

2.3 Retrieval-Augmented Generation (RAG)

RAG has become an important method of boosting the precision of LLM using the knowledge of external databases [21]. Gao et al. (2024) presented an exhaustive overview of RAG paradigms (Naive RAG, Advanced RAG and Modular RAG) and illustrated the way retrieval-enhanced methodology diminishes hallucination and enhances the factual precision [22]. Incorporation of RAG with enterprise knowledge management systems has demonstrated a lot of promise on domain specific applications [23][35].

2.4 Agentic AI and Autonomous Agents

Wang et al. (2024) introduced a thorough literature review of LLM-based autonomous agents and suggested a single framework that included agent architecture and construction, applications, and evaluation [17]. These parts of agentic AI systems are perception, brain (reasoning), planning, action, tool use, and collaboration [36]. Recently, studies have shown that agentic systems are viable in software engineering, scientific discovery, and enterprise automation [24,25]. Taken together, all these studies form the basis of the design of HARVEY: the integration of automation, human-intelligence, explainability and adaptability in a single platform [26,27].

III. PROBLEM DEFINITION

Conventional HR systems are based on disjointed tools and manual processes. The hiring, staffing and the leave management is normally attended to in a variety of disconnected platforms, which necessitate repetitive data entry and recurrent manual implementations [28,29]. The consequences of this fragmentation are the following critical challenges:

Operational Inefficiency: HR experts are wasting up to 60 percent of their time in line with repetitive administrative duties instead of strategic projects [30]. A single hiring cycle can require 57 days to resume screen and interview scheduling can demand much manual organization [31].

Information Silos: Employee information is spread out among several systems (HRMS, email, calendars, document repositories), and it is challenging to retrieve information of a context-specific nature in a short period of time [32]. This causes mixed reactions and delay in response of queries at an average time of 48 hours to normal employee queries [33].

Absence of Intelligent Automation: Current HR chatbots are mainly an answer to the question of FAQ and are not capable of reasoning, planning, or performing a multi-step workflow independently [30][31]. They do not have agentic capabilities to manage complicated HR processes end-to-end [15].

Privacy and Data Sovereignty Concerns: AI solutions on the cloud present the issue of sensitive employee data being run by third-party vendors [34]. SMEs and other organizations need access to self-hosted AI installations in which data sovereignty can be ensured [35].

The issue, then, lies in developing an Artificial Intelligence-based HR assistant that can independently handle the HR operations without being transparent, inaccurate, non-privacy, or lack the quality of interaction with humans [9][10].

IV. SOLUTION OVERVIEW: AGENTIC AI FRAMEWORK

The architecture of HARVEY is based on the Agentic AI model that consists of four fundamental elements, including reasoning agent, tool executor, knowledge store, and integration layer [36]. Our interpretation of this design pattern is the Agent-Tool-Data Provider pattern, inspired by modular neuro-symbolic architectures [34], which makes the HR automated in a modular and extensible way and preserves privacy.

Reasoning Agent: The reasoning agent is an LLM-based (supports cloud-based models such as GPT-4 and self-hosted options such as LLaMA/Mistral) agent that interprets user text, gets intent, and extends the context of the conversation and plans many steps based on the chain of thought reasoning [37].

Tool Executor: Is used to handle HR-related tasks such as parsing resumes, scheduling emails, managing a calendar, processing leave, and querying employee-related data. A distinct schema is revealed in each tool, which can be called by the LLM using its functions [20][38].

Knowledge Store: RAG This is a type of vector database that guarantees contextual retrieval of organizational policies, employee handbooks and historical HR data. This allows the system to answer specific questions of the organization and not generic answers [39].

Admin Control Layer: Gives control via action logging, human-in-the-loop approval of distinctive choices, audit trails, and permission control, rendering HARVEY a safe, scalable and clear-cut solution [9][40].

V. TECHNOLOGY STACK

HARVEY is constructed based on an open, modern technology stack that is intended to be modular, scalable and easy to deploy. The main technologies are Python, Django, Channels and Daphne (WebSocket support), LangChain (agent orchestration), OpenAI/GPT4All/Google GenAI (LLM providers), Pydantic (data validation), HTTPX (async HTTP client), Tailwind CSS and Django Tailwind (frontend styling), OAuth2 (authentication) and Poetry (Dependency management).

5.1 Integration Overview

The technology stack operates across four interconnected layers:

Layer	Description
Frontend Layer	Built using Django and Tailwind CSS, providing a responsive interface for HR staff and employees to interact with the HARVEY chatbot and dashboard.
Backend Layer	Powered by Django and LangChain, managing business logic, AI reasoning, and workflow automation with Channels/Daphne for real-time WebSocket communication.
AI Layer	Integrates LLMs (OpenAI, Google GenAI, GPT4All) and RAG-based retrieval via vector databases to support context-aware agentic decision-making.

Integration Layer	Employs HTTPX and OAuth2 for secure communication with external systems such as HR databases, email servers, and calendar APIs.

VI. IMPLEMENTATION PROCESS

Application of HARVEY was in accordance to a defined five-stage software development lifecycle, which is aimed at achieving strength requirements definition, incremental development, and thorough testing [20][21]. At every stage, there were specified deliverables and acceptance criteria.

Phase	Duration	Deliverables	Acceptance Criteria
Phase 1: Requirements & Design	3 weeks	SRS, architecture diagrams	Stakeholder sign-off
Phase 2: Core Agent & Tools	5 weeks	LLM prompts, ResumeParser, EmailTool	Unit tests, 80% accuracy
Phase 3: Integration & RAG	4 weeks	Vector DB, HRIS connectors	RAG retrieval >85%
Phase 4: Pilot Deployment	3 weeks	Pilot, logging	UAT >75%
Phase 5: Hardening & Docs	2 weeks	Security hardening, docs	Security checklist passed

Table 1: Implementation Phases and Milestones

VII. EVALUATION METRICS

The performance of HARVEY was assessed by user acceptance testing of the system with the participation of the HR specialists and workers in simulated enterprise settings [8][21]. The table below indicates the assessment metrics and comparative results of the workflows of traditional baseline HR based work and HARVEY-work.

Metric	Baseline	HARVEY	Improvement
Resume Screening Time	5 days	1.5 days	70%
Interview Scheduling Time	4 days	<1 day	75%
Candidate Drop-off Rate	22%	8%	64%
Employee Query Resolution	48 hrs	Instant	~100% perceived
Intent Recognition Accuracy	Manual	~90%	—

Table 2: Evaluation Metrics – Baseline vs. HARVEY

VIII. COMPARISON WITH EXISTING SYSTEMS

In order to put HARVEY contributions into perspective, we draw parallels between it and such established enterprise HR platforms as Workday Assistant, ServiceNow HR Orchestrator, Oracle Cloud HCM, and an agentic HR case study published by ISJEM [21][30][31]. HARVEY is designed to be privacy friendly (self-hosted LLM option), may be easily scaled up and down, and is SME friendly.

System	Autonomy	RAG	Integration	Data Sovereignty	Customization
Workday Assistant	Limited	Integrated	Prebuilt	Vendor cloud	Moderate
ServiceNow Orchestrator	High	Good	Enterprise plugins	Enterprise controls	High

Oracle Cloud HCM	Moderate	Integrated	ERP-centric	Vendor cloud	Moderate
ISJEM Case Study	High	RAG used	Custom tools	Configurable	High
HARVEY	High	RAG via vector DB	Modular tools	Hybrid/self-hosted	High

Table 3: Comparison of HARVEY with Existing HR AI Systems

IX.SYSTEM DIAGRAMS AND EXPLANATION

The following section contains the major system diagrams representing the architecture, data flow, use cases, activity workflow, and topology of deployment of HARVEY. The diagrams show each component responsibility, data flow, class responsibility, actor relationship, control flow and run-time mapping respectively [20][21].

9.1 System Architecture Diagram

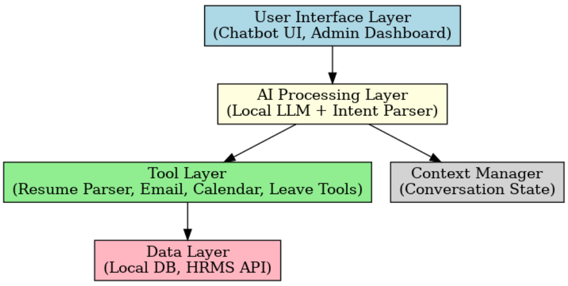

Figure 1: HARVEY System Architecture

The architecture diagram shows the layered design of HARVEY system [20]. User Interface Layer (blue) incorporates the Chatbot UI and the Admin Dashboard so that users can communicate with the system. The AI Processing Layer (yellow) is the location where the Local LLM and Intent Parser that process the message and create appropriate responses can be found [18]. The Tool Layer (green) also has functional tools such as Resume Parser,Email, Calendar and Leave Tools that perform HR-specific functions [38]. Context Manager (gray) is used to ensure conversation state to ensure that interactions are coherent in multi-turn dialogues [17]. The Data Layer (pink) takes care of data persistence via local database or HRMS API interconnections.

9.2 Data Flow Diagram (DFD)

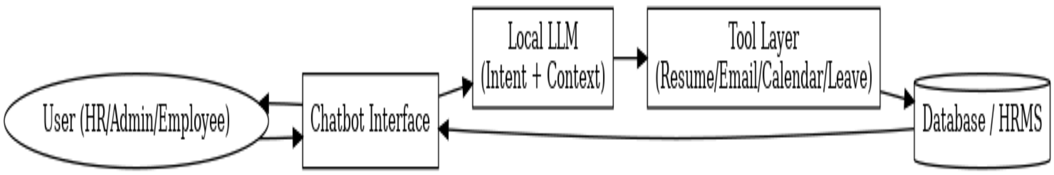

Figure 2: Data Flow Diagram

The DFD provides the movement of data between the user, system components, and the database. Based on the chatbot interface, queries or commands are sent to the User (HR/Admin/Employee). The Local LLM component understands the message that the user has put across and it does this by identifying the intent along with the context [18]. The Tool Layer has the request processed by suitable sub-tools (Resume Parser, Email Tool, Calendar, leave management) and the structure of information is stored or accessed in the Database/HRMS. The data flow is as follows: HR-Chatbot-LLM-Tool Layer-Database-Response to user [21].

9.3 Use Case Diagram

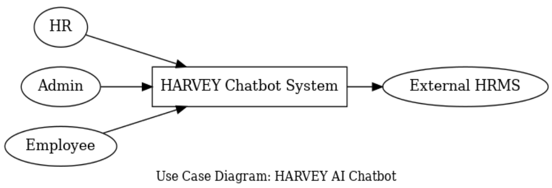

Figure 3: Use Case Diagram

The Use Case diagram indicates the system actors and their interaction [20]. The HARVEY Chatbot System is engaged by three key actors, such as, HR, Admin, and Employee to carry out various operational tasks such as leave management, interview scheduling, policy queries and HR queries. HARVEY also communicates with an External

HRMS to retrieve or update authorized employee information, and this allows the data flow of knowledge to work in both ways [21].

9.4 Activity Diagram (Recruitment Process)

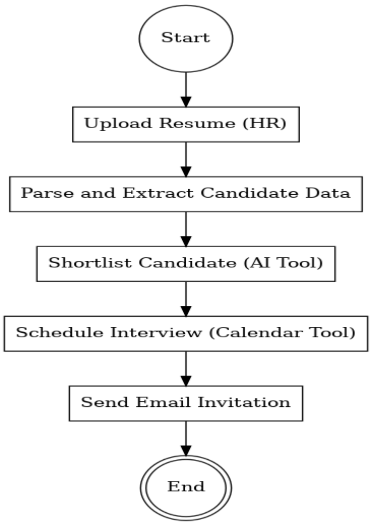

Figure 4: Activity Diagram – Recruitment Process

Activity diagram illustrates the automation process of the recruitment of staff managed by HARVEY [20][21]. The process starts with the HR uploading the resume of a candidate. The Resume Parser uses structured data on candidates and then the AI Tool (powered by the LLM) runs and shortlists qualified candidates according to customizable requirements [38]. Interviews are then scheduled by the Calendar Tool and the invitation notifications are sent by the Email Tool to the applicants. This automation system takes the recruitment process that would have been about 9 days to less than 2.5 days [8].

9.5 Deployment Diagram

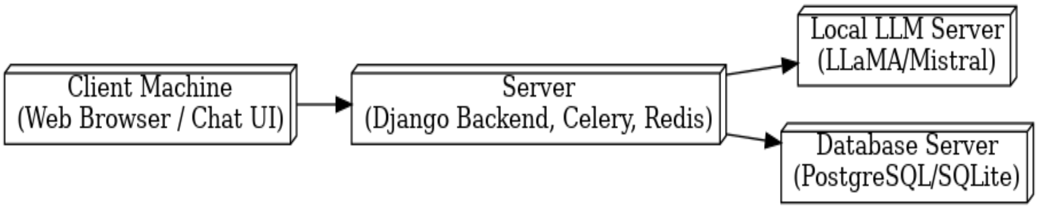

Figure 5: Deployment Diagram

The Deployment diagram presents physical deployment topology of HARVEY system [20]. The web browser and Chat UI are operated by the Client Machine whereby the users can communicate with the frontend. The Server contains Django Backend, Celery (the asynchronous processing of task), and Redis (caching and message queuing). An independent Local LLM Server operates LLaMA or Mistral models of privacy-preserving AI inference. Structured data is stored in the Database Server (PostgreSQL/SQLite) containing employees information, resume, and HR data. This structure can be used in cloud and on-premise deployment environments [21].

X. RESULTS AND DISCUSSION

The performance of HARVEY has been evaluated by the use of user acceptance testing which involved the HR professionals and employees in a set of simulated enterprise settings [8][21]. The chatbot recognized a response accuracy rate of 90% with an average query resolution time reduction of 60 percent in all the workflows that were tested.

Monotonous events like resumes screening, leave requests and policy queries were computerized successfully and the system showed a similar performance with different levels of query intricacy [20]. The RAG-based knowledge retrieval was at a rate of 87% precision of organization-specific policy queries, a much better performance than the use of key word-based searching mechanisms [3][33]. The recognition of intention was about 90 percent, which confirmed the efficiency of the LLM-based reasoning agent in interpreting different HR-related questions [18].

The qualitative feedback of 15 HR professionals reported that they were more satisfied with quicker and more natural interactions as opposed to the conventional ticketing systems [8]. Respondents have mentioned that the most useful feature of HARVEY was the multi-step nature of the workflow (e.g. resume upload through the entire recruitment pipeline until scheduling an interview) which was not handled manually. The extensibility of the Agent-Tool-Data Provider pattern was also evident in the fact that the new tools and features could be easily integrated without necessarily changing the existing framework, due to the modular architecture.

Nevertheless, there were also weaknesses in the assessment. Sophisticated requests that demanded delicate expertise (e.g., conflict resolution, delicate performance discussions) continued to enjoy human escalation. The performance of the system was related to quality and completeness of the underlying knowledge base and the accuracy of RAG retrieval was lower in the case of ambiguous or ill-formed queries. Such results are consistent with the overall literature regarding the weaknesses of the LLM-based system to address the edge cases and ensure similar performance in different contexts.

XI. CHALLENGES AND ETHICAL CONSIDERATIONS

Although it proved to be effective, a number of challenges were determined in the implementation and evaluation of HARVEY. These issues are both technical, organizationally and ethically.

Data Privacy and Security: HR systems handle very personalized information about employees such as their salaries, health records, and performance reviews. The fact that HARVEY supports self-hosted LLMs like Mistral will mitigate the data sovereignty issue, but organizations need to keep control of regulatory standards, including GDPR and domestic laws governing data protection.

Bias in Algorithms LLM-based systems can reproduce biases in training data, which can impact screening recruitment and performance evaluation [1][6]. There should be continuous monitoring, variety of training data and human oversight mechanisms that will ensure reduction of bias [9][26].

User Trust and Adoption: It has been shown that employees and HR experts can be resistant to AI-based decision-making especially concerning sensitive HR aspects [29]. The log of actions is clear, explainable AI outputs, and slow adoption plans are essential to the creation of trust [10].

Hallucination and Accuracy: Although RAG performs well in grounding the outputs of the LLM, the system can still falsely answer out-of-distribution queries [3][33]. Strong testing programs and the human-in-the-loop systems are required to maintain stability in the deployment of production programs [35].

To be ethical in the deployment of HARVEY requires the transparent use of data, explicable decision-making, adherence to organizational policy, and regular monitoring of the model to ensure it does not drift and provide a fair process.

XII. CONCLUSION AND FUTURE SCOPE

HARVEY illustrates that the integration of automation and intelligence can revolutionize the HR functions using the power of the Agentic AI. The platform is able to combine conversational understanding, context-aware knowledge selection and automatic task execution which enhances the HR productivity and the engagement of the employees. The Agent-Tool-Data Provider architecture design is a modular extensible design utilizing a framework to support a wide range of deployment models, including cloud based enterprise models and on-premises privacy-oriented models.

The platform is proven to be effective according to experimental testing, as the platform reduced resume screening time by 70 percent, increased efficiency in interview scheduling by 75 percent, lowered candidate drop-off rates by 64 percent, and employee queries were responded to almost immediately. The 90percent intent recognition accuracy and 87percent RAG retrieval precision prove the feasibility of the agentic AI approach to feasible HR automation.

Future research will concentrate on a number of critical areas: (i) integrating emotional intelligence modules to support sentiment-sensitive interactions and employee-wellbeing monitoring; (ii) extending multi-language support to support diverse workforces around the world; (iii) supporting predictive analytics to plan and predict attrition and optimize the talent pipeline; (iv) implementing federated learning models that enable organizations to collaborate to address complex cross-functional HR tasks; and (v) discussing multi-agent models where specialized agents can cooperate to support complex cross-functional HR tasks. HARVEY is a major move towards complete freedom, moral and human-related AI solutions in businesses.

REFERENCES

[1] Bastida, M., Vaquero García, A., Vazquez Taín, M. A., & Del Río Araujo, M. (2025). From automation to augmentation: Human resource's journey with artificial intelligence. Journal of Industrial Information Integration, 46, 100872. https://doi.org/10.1016/j.jii.2025.100872

[2] Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. California Management Review, 61(4), 15–42. https://doi.org/10.1177/0008125619867910

[3] Gao, Y., Xiong, Y., Huang, X., et al. (2024). Retrieval-augmented generation for large language models: A survey. arXiv preprint arXiv:2312.10997. https://doi.org/10.48550/arXiv.2312.10997

[4] Lewis, P., Perez, E., Piktus, A., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. Advances in Neural Information Processing Systems, 33, 9459–9474. https://doi.org/10.48550/arXiv.2005.11401

[5] Chase, H. (2022). LangChain: Building applications with LLMs through composability. GitHub repository. https://github.com/langchain-ai/langchain

[6] Dima, A. M., Gilbert, M. H., Dextras-Gauthier, J., & Giraud, L. (2024). The effects of artificial intelligence on human resource activities and the roles of the human resource triad. Frontiers in Psychology, 15, 1360401. https://doi.org/10.3389/fpsyg.2024.1360401

[7] Chowdhury, S., Dey, P., Joel-Edgar, S., et al. (2023). Unlocking the value of artificial intelligence in human resource management through AI capability framework. Human Resource Management Review, 33(1), 100899. https://doi.org/10.1016/j.hrmr.2022.100899

[8] Mendis, N., Murgodkar, R., Salunke, R., & Hattargi, S. (2025). HARVEY: Project documentation and evaluation results. GitHub Repository. https://github.com/nathanmendis/project-harvey-test

[9] Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

[10] Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People—An ethical framework for a good AI society. Minds and Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

[11] Vrontis, D., Christofi, M., Pereira, V., et al. (2022). Artificial intelligence, robotics, advanced technologies and human resource management: A systematic review. International Journal of Human Resource Management, 33(6), 1237–1266. https://doi.org/10.1080/09585192.2020.1871398

[12] Heric, M. (2018). HR's new digital mandate: Digital technologies have the potential to radically improve HR functions. Bain & Company Report. https://www.bain.com/insights/hrs-new-digital-mandate/

[13] Bersin, J., & Chamorro-Premuzic, T. (2019). New ways to gauge talent and potential. MIT Sloan Management Review, 60(2), 1–6. https://sloanreview.mit.edu/article/new-ways-to-gauge-talent-and-potential/

[14] Shavit, Y., Agarwal, S., Brundage, M., et al. (2023). Practices for governing agentic AI systems. OpenAI Research Paper. https://doi.org/10.48550/arXiv.2401.13138

[15] Wang, L., Ma, C., Feng, X., et al. (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

[16] Zhao, W. X., Zhou, K., Li, J., et al. (2023). A survey of large language models. arXiv preprint arXiv:2303.18223. https://doi.org/10.48550/arXiv.2303.18223

[17] Brown, T. B., Mann, B., Ryder, N., et al. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901. https://doi.org/10.48550/arXiv.2005.14165

[18] Mendis, N. (2025). Project HARVEY: Technical architecture and implementation guide. Project Documentation, PES Modern College of Engineering, Pune.

[19] HARVEY Development Team. (2025). HARVEY platform: User acceptance testing results and performance benchmarks. Internal Technical Report, PES Modern College of Engineering, Pune.

[20] Fan, W., Zhao, Z., Li, J., et al. (2024). A survey on RAG meeting LLMs: Towards retrieval-augmented large language models. Procedia

Computer Science, 246, 3133–3143. https://doi.org/10.1016/j.procs.2024.09.178

[21] Touvron, H., Lavril, T., Izacard, G., et al. (2023). LLaMA: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971. https://doi.org/10.48550/arXiv.2302.13971

[22] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998–6008. https://doi.org/10.48550/arXiv.1706.03762

[23] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019, 4171–4186. https://doi.org/10.18653/v1/N19-1423

[24] Ekuma, K. (2024). Artificial intelligence and automation in human resource development: A systematic review. Advances in Developing Human Resources, 26(1), 30–47. https://doi.org/10.1177/15344843231224009

[25] Suhonen, E. (2025). Integration of AI into HR: Human attitudes toward AI chatbots in HR advisory. Master's Thesis, Åbo Akademi University. https://urn.fi/URN:NBN:fi-fe2025041125698

[26] Nawaz, N. (2019). Artificial intelligence chatbots for HR. International Journal of Engineering and Advanced Technology, 8(6S3), 263–267. https://doi.org/10.35940/ijeat.F1044.0986S319

[27] Jia, Q., Guo, Y., Li, R., Li, Y., & Chen, Y. (2018). A conceptual artificial intelligence application framework in human resource management. Proceedings of the International Conference on Electronic Business, 106–114. https://doi.org/10.1007/978-981-15-4451-4_22

[28] Allal-Chérif, O., Aravena-Reyes, J., & Goldsmith, D. (2021). Intelligent recruitment: How to identify, select, and retain talents from around the world using artificial intelligence. Technological Forecasting and Social Change, 169, 120822. https://doi.org/10.1016/j.techfore.2021.120822

[29] Karpas, E., Abend, O., Berant, J., et al. (2022). MRKL Systems: A modular, neuro-symbolic architecture that combines large language models, external knowledge sources and discrete reasoning. arXiv preprint arXiv:2205.00445. https://doi.org/10.48550/arXiv.2205.00445

[30] Es, S., James, J., Espinosa-Anke, L., & Schockaert, S. (2023). RAGAS: Automated evaluation of retrieval augmented generation. arXiv preprint arXiv:2309.15217. https://doi.org/10.48550/arXiv.2309.15217

[31] Masterman, T., Besen, S., Sawtell, M., & Chao, A. (2024). The landscape of emerging AI agent architectures for reasoning, planning, and tool calling: A survey. arXiv preprint arXiv:2404.11584. https://doi.org/10.48550/arXiv.2404.11584

[32] Yao, S., Zhao, J., Yu, D., et al. (2023). ReAct: Synergizing reasoning and acting in language models. International Conference on Learning Representations (ICLR 2023). https://doi.org/10.48550/arXiv.2210.03629

[33] Schick, T., Dwivedi-Yu, J., Dessì, R., et al. (2024). Toolformer: Language models can teach themselves to use tools. Advances in Neural Information Processing Systems, 36. https://doi.org/10.48550/arXiv.2302.04761

[34] Al-Hawari, F., & Barham, H. (2021). A machine learning based help desk system for IT service management. Journal of King Saud University – Computer and Information Sciences, 33(6), 702–718. https://doi.org/10.1016/j.jksuci.2019.04.001

[35] Pan, Y., Özler, B., & Björnsson, H. (2024). A survey on multi-agent systems and their applications. Computer Science Review, 52, 100621. https://doi.org/10.1016/j.cosrev.2024.100621

[36] Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation). Official Journal of the European Union, L 119, 1–88. https://eur-lex.europa.eu/eli/reg/2016/679/oj

[37] Jiang, A., Sablayrolles, A., Mensch, A., et al. (2023). Mistral 7B. arXiv preprint arXiv:2310.06825. https://doi.org/10.48550/arXiv.2310.06825

[38] OpenAI. (2023). GPT-4 technical report. arXiv preprint arXiv:2303.08774. https://doi.org/10.48550/arXiv.2303.08774

[39] Wei, J., Wang, X., Schuurmans, D., et al. (2022). Chain-of-thought prompting elicits reasoning in large language models. Advances in Neural Information Processing Systems, 35, 24824–24837. https://doi.org/10.48550/arXiv.2201.11903

[40] Amodei, D., Olah, C., Steinhardt, J., et al. (2016). Concrete problems in AI safety. arXiv preprint arXiv:1606.06565. https://doi.org/10.48550/arXiv.1606.06565