Evaluating Generative AI Technology Choices and Software Frameworks for Developing AI Solutions in Business

Umair Ali Khan[1], Dmitry Kudryavtsev[1], Janne Kauttonen[1], Antti Joutsenniemi[1], Akseli Leskinen[1], Jukka Remes[1], Roman Yangarber[2], Lidia Pivovarova[2] and Yiheng Wu[2]
[1]Digital Transition and AI, Haaga-Helia University of Applied Sciences, Helsinki
[2]Department of Digital Humanities, University of Helsinki, Helsinki

UmairAli.Khan@haaga-helia.fi
dmitry.kudryavtsev@haaga-helia.fi
Janne.Kauttonen@haaga-helia.fi
antti.joutsenniemi@haaga-helia.fi
akseli.leskinen@haaga-helia.fi
jukka.remes@haaga-helia.fi
Roman.Yangarber@helsinki.fi
lidia.pivovarova@helsinki.fi
yiheng.wu@helsinki.fi

Abstract: Generative AI (GenAI) is being increasingly adopted in business, yet leveraging it effectively remains challenging due to diverse use cases and organizational needs. While generic end-user tools are widely used, they often fall short in addressing domain-specific tasks and integration requirements. This paper reviews the limitations of such tools and emphasizes the need for customized, use-case-specific GenAI solutions. We examine emerging software frameworks for developing tailored GenAI solutions.

Keywords: Generative AI, Technology choices, Generative AI frameworks, Enterprise AI integration, Generic and custom GenAI solutions

1. Introduction

GenAI adoption is increasing worldwide. According to a global survey (McKinsey, 2025), nearly 75% of organizations now use AI in at least one function. However, realizing long-term value remains challenging, particularly as organizations struggle to choose the right tools amid a growing landscape of frameworks and technologies (Deloitte, 2025). Most are still in the experimentation phase, using proof of concepts (PoCs) to test feasibility.

This paper reviews the current state of GenAI adoption, examining the strengths and limitations of popular end-user tools, and the growing need for custom solutions. It also reviews key development frameworks.

2. Generative AI Adoption in Business Context

Our AI consultancy experience in the Finnish AI Region European Digital Innovation Hub (FAIR EDIH) (Finnish AI Region, n.d.) project reveals that both large and small companies belonging to different sectors are adopting AI. The major sectors include healthcare, education, construction, manufacturing, real estate, and logistics, among others (Khan et al., 2025). Businesses are increasingly turning to GenAI to boost productivity, automate labor-intensive tasks, such as content creation and document drafting, reduce operational costs, and gain a competitive advantage (Deloitte, 2025; McKinsey, 2023).

Most companies start their GenAI journey with popular end-user tools such as ChatGPT, MS Copilot, NotebookLM, and Perplexity AI. Microsoft Copilot is integrated with Microsoft 365 applications and supports tasks like email drafting, document summarization, and generating data insights. However, it offers limited flexibility outside the Microsoft ecosystem, lacks transparency, and presents integration challenges with non-Microsoft platforms. In our AI consultancy, several companies have also shown dissatisfaction with its performance for their tailored tasks.

NotebookLM provides summarization, citation generation, and audio guides from uploaded content. However, it struggles with poorly formatted PDFs, limited file support, and strict usage limits. Despite assurances of data privacy, concerns remain regarding content handling.

Conversational agents such as ChatGPT, Gemini, Claude, and Mistral offer natural language interaction with support for various input types and basic web integration. Their ease of use and wide accessibility are clear advantages. However, they lack organization-specific knowledge, are prone to hallucinations, and cannot be directly embedded into business workflows.

Search engines like Perplexity AI combine web search and AI reasoning, offering real-time, sourced answers with file upload and API options. They excel at information delivery but depend heavily on the quality of external sources and prioritize information retrieval over creative or specialized output. Like other tools, it also raises privacy concerns.

3. Need for Customized GenAI Solutions

Enterprises are increasingly recognizing that one-size-fits-all generative AI tools often fall short for specialized needs and company-specific use-cases. The generic AI solutions suffice for common and recurring cases such as simple content generation, basic data analysis, finding information from internal documents with simple queries, quick internet searches, and simple knowledge extraction. The issues of hallucination and inaccurate answers, lack of transparency, minimal control over the implementation pipelines and model behavior, lack of use case adaptability, lack of data privacy, and lack of integration in organizational workflows warrant the development of custom AI solutions.

During our AI consultancy (Finnish AI Region, n.d.), we encountered several use cases that entail the development of custom AI tools. Some of these use-cases include generation of sales quotations from multiple documents with multiple data formats, combining tacit and documented knowledge in AI assistants, dynamic generation of programming codes based on new use-cases, automatic resolution of customer complaints, matchmaking and networking, and the AI systems that improve their knowledge with user interactions, to name a few.

Generic end-user tools operate as standalone assistants and cannot directly integrate with an enterprise's IT systems or workflows. For instance, ChatGPT cannot automatically retrieve real-time internal data or trigger actions in business applications. In contrast, custom GenAI can be built into products and processes, connecting to CRM systems, knowledge bases, ERPs, or any relevant APIs (Van der Heijden, 2023). In particular, new advancements in GenAI that enable LLMs to connect to any data source, such as Anthropic's Model Context Protocol (MCP) (Anthropic, 2024), open new areas of custom AI solutions. Table I presents a comparison of generic and custom AI tools.

Table 1: Comparison of generic and custom AI tools

GenAI Tool Type	Use-Case Fit	Data Privacy	Scalability & Cost	Customization	Regulatory Compliance	Transparency & Control
Generic Tools	Limited	Conditional (e.g., MS Copilot)	Limited (costly at scale)	None	Poor	Low (black box)
Custom Tools	High	High (open-source models or secure cloud)	Flexible (tailored infra)	High (fine-tunable)	High	Full (pipeline control)

4. Software Frameworks for Developing GenAI Solutions

Developing GenAI solutions requires choosing software frameworks that align with the use case, organizational goals, and technical capacity. Code-based frameworks offer high flexibility and control, making them ideal for complex, custom applications, but they demand technical expertise and longer development time. Low/no-code platforms enable faster deployment with minimal effort, making them suitable for prototyping and simpler use cases, though they offer limited customization and may lead to vendor lock-in. Table II summarizes some low/no-code software frameworks for developing AI solutions.

Table 2: A comparison of popular low/no-code software frameworks suitable for GenAI solution development

Framework	Type & Features	Ease of Use	Common Use Cases	Ecosystem	Scalability	API/Integration
n8n (n8n, 2025)	Visual, open-source workflow builder	Moderate	Automation, custom bots	Large (400+ integrations)	Cluster-based	Webhooks, REST, custom nodes

Framework	Type & Features	Ease of Use	Common Use Cases	Ecosystem	Scalability	API/Integration
Motia (Motia, 2025)	Code-first, event-driven agents	Steep	Custom AI automation	Small, growing	Modular, cloud-ready	Extendable HTTP APIs
Flowise (Flowise, 2025)	No-code LLM builder (LangChainJS)	Easy	RAG apps, chatbots	Active open-source	Self-hosted/cloud	Custom nodes, APIs
Dify (Dify, 2025)	Low-code LLM with ops tools	Low	Internal tools, SaaS AI	Popular on GitHub	Kubernetes-ready	REST APIs, plugin support
Botpress (Botpress, 2025)	Conversational AI, modular NLU	Shallow	Multi-channel bots	Mature, large base	Horizontal	SDKs, API hooks
Langflow (Langflow, 2025)	Drag-and-drop LangChain canvas	Easy	Q&A bots, prototyping	Emerging open-source	Limited (SMB scale)	Python components, API access

LLMs can be customized in several ways. Fine-tuning is a primary method for customizing LLMs by further training them on task- or domain-specific datasets to improve performance on specialized tasks. This approach has been useful in adapting LLMs to various applications. For instance, fine-tuning has been applied to enhance domain-specific machine translation capabilities, addressing the limitations of general-purpose models in specialized fields like legal and medical translations (Hu et al., 2025).

The other way to customize LLM is Retrieval Augment Generation, which augments an LLM with an external information retrieval component to overcome the model's knowledge limitations (such as fixed training data or hallucinations). In this setup, the system fetches relevant documents or facts from a knowledge base in response to a query, and the LLM generates its output conditioned on that retrieved context (Li et al., 2024).

Integrating LLMs with knowledge graphs (KGs) is also an emerging technique to integrate structured, verified knowledge into generation processes. KGs store facts as interconnected entities and relations, offering a reliable and interpretable source of truth that can ground an LLM's responses. By incorporating KG data, the LLM's outputs can become more consistent and factual, helping to reduce hallucinations in complex reasoning tasks (Ji et al., 2024).

In more advanced systems, AI agents are designed to autonomously plan and execute tasks, enabling GenAI systems to achieve goals with minimal human input (Deloitte, 2025). Agentic frameworks allow these agents to dynamically interact with external data, APIs, and tools, supporting multi-step workflows, automation, and context-aware responses. Potential applications include report generation and information retrieval and validation (e.g., GPT Researcher (GPTR, 2025)). Table III compares some agentic frameworks.

RAG and agentic AI solutions can be developed using both code-based and low/no-code software frameworks for developing AI solutions. RAG improves response accuracy and mitigates hallucinations by retrieving relevant information from external sources before generating output. Frameworks like LangChain (LangChain, 2025) and LlamaIndex (LlamaIndex, 2025) enable RAG development by integrating LLMs with internal data or knowledge bases to power chatbots, recommendation engines, matchmaking tools, and document-based knowledge extraction.

Table3: A comparison of popular agentic frameworks for developing AI solutions

Software Framework	Key Features	Ease of Use	Use Cases	Ecosystem	Scalability	API Support	Limitations
LangGraph (LangGrah, 2025)	Visual, graph-based orchestration (LangChain-integrated)	Moderate	Enterprise AI apps (e.g., chatbots, CX)	Large, growing	High (prod-ready)	Agent workflow APIs	Limited docs, partial cloud support
BeeAI (BeeAI, 2025)	No/low-code, multi-agent	Easy	IT ops, internal AI tools	Small, enterprise-backed	High (parallel tasks)	Python/TS SDKs	Small ecosystem, maturing APIs

Software Framework	Key Features	Ease of Use	Use Cases	Ecosystem	Scalability	API Support	Limitations
	orchestration						
AutoGen (AutoGen, 2025)	Modular, Microsoft-backed, dynamic workflows	Moderate	Coding agents, supply chain, multi-agent ops	Strong research/enterprise	Very high (event-driven)	Human-in-loop APIs	Python-only, needs infra setup
CrewAI (CrewAI, 2025)	Lightweight, open-source, quick setup	Very easy	SMB workflows, content creation	Active grassroots	Cloud-ready (AWS/Azure)	Simple APIs	No streaming, limited model support

Anthropic's MCP is an open standard that connects LLMs to diverse data sources, enabling secure, flexible, and interoperable agent-based applications through a client-server architecture (Andreessen, 2025). It holds potential to enhance integration and data security across enterprise systems.

Selecting the right LLM is equally important for custom AI development. Proprietary LLMs (such as GPT-4o, O3, Claude-3.7-sonnet, and Gemini 2.5 Pro) offer strong capabilities but may lead to vendor dependency. Open-source alternatives (such as DeepSeek-V3, LLaMA 3, Qwen2.5, and Mistral Large Instruct) provide greater control for in-house use, though they require robust (hardware-software) infrastructure. A hybrid approach starts with proprietary models in secure environments (e.g., AWS, Azure) and transitions to open-source models. Key selection criteria include accuracy, context length, multimodality, latency, compute needs, token pricing, and API/tooling ecosystem.

A summary of the technologies for creating GenAI solutions is presented in Fig. 1.

5. Conclusion

This study highlights the limitations of generic GenAI tools in meeting the domain-specific and integration needs of businesses, justifying the growing necessity for custom-built solutions. While commercial platforms offer accessibility and a quick start, they often lack the transparency, control, and adaptability required for enterprise applications. Custom GenAI systems, integrated with organizational workflows and data environments, are better positioned to deliver lasting value and ensure regulatory compliance.

Effective AI adoption further demands strategic alignment with business goals, a thoughtful evaluation of frameworks, and a realistic assessment of integration complexity and internal expertise. A phased roadmap is essential: starting with commercial tool exploration, identifying organizational gaps, and progressing toward tailored solutions through co-development with technical experts. Building AI literacy across roles and collaborating with research and innovation networks further strengthens implementation.

With the identification of these gaps and needs, we are developing a GenAI toolkit for knowledge management in business that will allow businesses to create their custom, use-case-specific knowledge management applications.

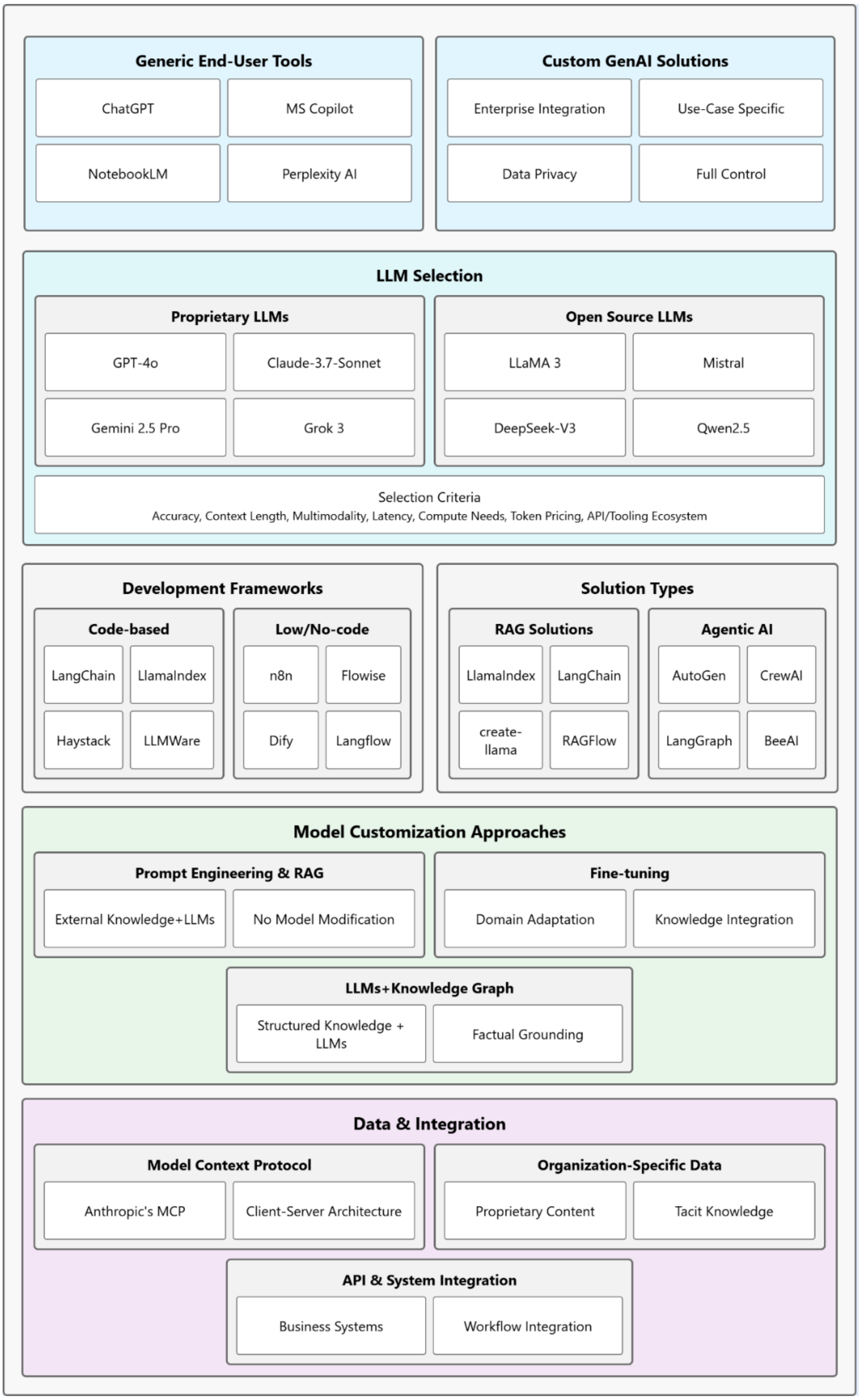

Figure 1: Overview of technologies for creating GenAI solutions

Acknowledgements

The present work is part of the Generative AI-Enhanced Knowledge Management in Business (GAIK) project, which is co-funded by the European Union, European Regional Development Fund (Euroopan Aluekehitys rahasto, EAKR).

Ethics Declaration: No ethical clearance was required for this research.

AI Declaration: An AI tool was used for rephrasing and summarizing some parts of the paper with careful validation.

References

Andreessen Horowitz. (2025). *A deep dive into MCP and the future of AI tooling*. Retrieved from https://a16z.com/a-deep-dive-into-mcp-and-the-future-of-ai-tooling/

Anthropic. (2024). Model Context Protocol. https://www.anthropic.com/news/model-context-protocol

BeeAI. (2025). BeeAI.dev. https://beeai.dev/

Botpress. (2025). Botpress. https://botpress.com/

CrewAI. (2025). CrewAI [AI agent orchestration platform]. https://www.crewai.com/

Deloitte AI Institute. (2025). *The state of generative AI in the enterprise*. Deloitte. https://www2.deloitte.com/us/en/pages/consulting/articles/state-of-generative-ai-in-enterprise.html

Dify. (2025). Dify.ai. https://dify.ai/

Finnish AI Region. (n.d.). https://www.fairedih.fi/

FlowiseAI. (2025). FlowiseAI. https://flowiseai.com/

GPTR. (2025). GPTR.dev. https://gptr.dev/

Hu, Yinghao, et al. (2025) *Fine-tuning Large Language Models for Improving Factuality in Legal Question Answering*. arXiv preprint arXiv:2501.06521.

Ji, Yixin, et al. (2024) *Retrieval and reasoning on KGs: Integrate knowledge graphs into large language models for complex question answering*. Findings of the Association for Computational Linguistics: EMNLP 2024. 2024.

Khan et al. "AI Adoption in Finnish SMEs: Key Findings from AI Consultancy at a European Digital Innovation Hub." *IEEE 23rd World Symposium on Applied Machine Intelligence and Informatics*. 2025.

LangChain. (2025). LangChain. https://www.langchain.com/

Langflow. (2025). Langflow. https://www.langflow.org/

LangGraph. (2025). LangGraph. https://www.langchain.com/langgraph

Li, Zhuowan, et al. (2024) *Retrieval augmented generation or long-context LLMs? a comprehensive study and hybrid approach*. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track.

LlamaIndex. (2025). LlamaIndex. https://www.llamaindex.ai/

McKinsey & Company. (2023). *The state of AI in 2023: Generative AI's breakout year*. Retrieved from https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-yea

McKinsey & Company. (2025, March). *The state of AI: How organizations are rewiring to capture value*. https://www.mckinsey.com/~/media/mckinsey/business%20functions/quantumblack/our%20insights/the%20state%20of%20ai/2025/the-state-of-ai-how-organizations-are-rewiring-to-capture-value_final.pdf

Microsoft. (2025). AutoGen. https://microsoft.github.io/autogen/stable/

Motia. (2025). Motia.dev. https://motia.dev/

n8n. (2025). n8n. https://n8n.io/

Van der Heijden T. (2023). *How domain-specific generative AI will shape the future for enterprise ESG teams*. Retrieved from https://www.briink.com/post/how-domain-specific-generative-ai-will-shape-the-future-for-enterprise-esg-teams