

How to Make Informed Choices When Deploying Generative AI in HR

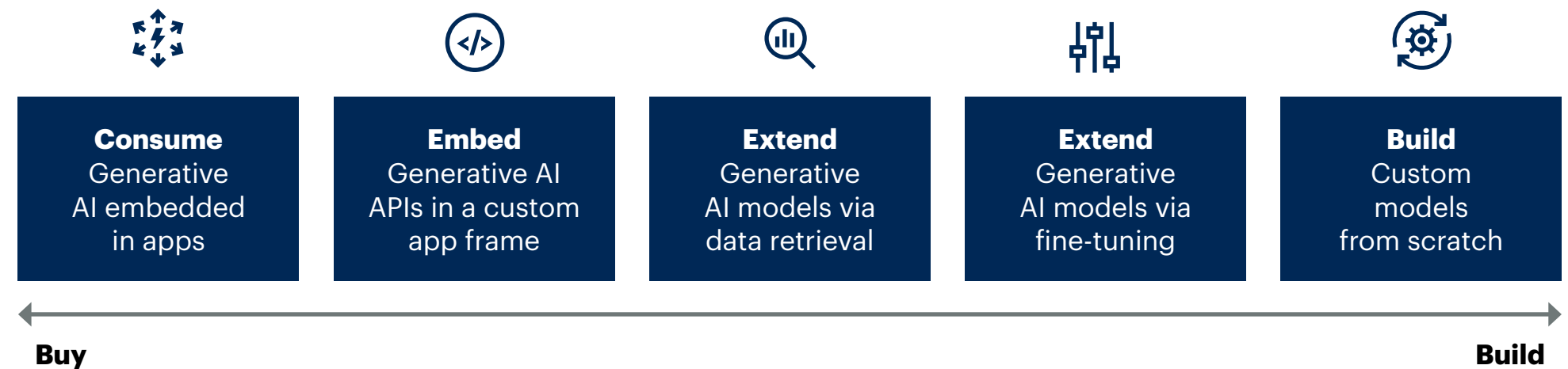
by Hiten Sheth, Leinar Ramos, Eser Rizaoglu and Helen Poitevin

HR leaders seeking to benefit from generative AI face a confusing array of options that require collaboration with IT partners and executives. Here, we compare various deployment approaches and offer guidelines to make informed choices amid rising investments by HR tech providers.

The proliferation of generative AI (GenAI) applications, such as ChatGPT, has led to heightened interest in GenAI within the industry, presenting HR leaders with new opportunities and challenges. Gartner's benchmark poll data highlights that 81% of HR leaders have implemented or are exploring GenAI solutions.¹ Recent Gartner interactions with HR leaders reveal many are considering a range of GenAI deployment tactics, from consumer-facing GenAI tools to the formation of in-house AI/ML teams. However, many are also unaware of the different approaches to deploying GenAI and their varying complexities. Furthermore, an HR leader's failure to consider risks and strategize for long-term success can impede GenAI adoption within the HR function.

While most HR leaders may predominantly consume GenAI through their existing HR vendor platforms, they must also familiarize themselves with alternative methods of access. This broad understanding empowers HR leaders to communicate and collaborate more effectively with their IT partners and strengthens their ability to make well-informed choices.

» Figure 1: Five Key Approaches to Generative AI Deployment



Source: Gartner

By engaging with IT and vendor partners, HR leaders ensure the GenAI tools they choose align with HR's unique needs and use cases. This report provides guidance on various ways of accessing GenAI and highlights necessary partnerships and considerations for thoughtful GenAI adoption.

GenAI Deployment Approaches

Gartner sees five key emerging approaches to consuming generative AI capabilities (see Figure 1). For insights into how these approaches translate to practical applications, explore the GenAI use cases.

Consume

The “consume” approach is the simplest way to utilize GenAI capability within the HR function. This approach is like turning on your existing HR technology providers’ out-of-the-box generative AI capabilities.

In this case, HR needs to track its vendors’ roadmap and adopt functionality as it emerges. HR leaders must evaluate vendors’ GenAI capabilities to ensure use cases add value and adhere to compliance requirements.

Embed

The “embed” approach to GenAI deployment involves connecting to external GenAI tools or services from within your HR applications. It makes those capabilities accessible to users within a widget, window or something similar in the HR applications that allow such extensions. This approach is advantageous in cases where your HR applications do not include vendor-provided GenAI features or the vendor does not plan to introduce them in their future roadmap.

HR leaders must collaborate closely with internal IT and technology architecture teams to ensure alignment with their organization’s data, security and compliance standards. This approach may have limitations in driving domain-specific use cases, as external GenAI models are typically not trained on specific HR domains. The nature and quality of the prompts will impact output quality so you will need staff with prompt engineering skills to realize the value of this approach.

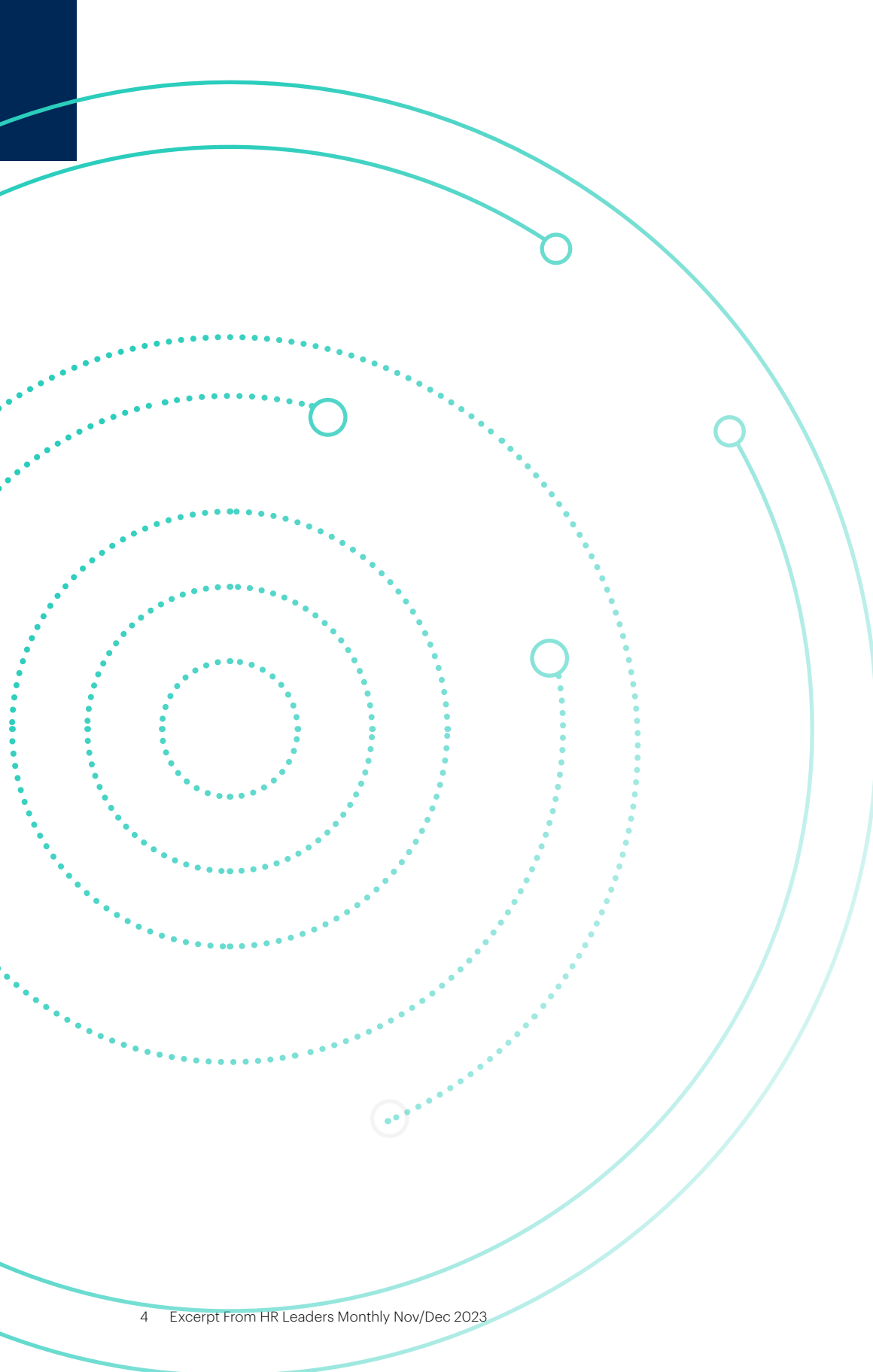
Extend via Data Retrieval

Within the HR function, the “extend via data retrieval” approach can improve the accuracy and quality of GenAI model responses for HR-specific tasks. This approach, known as retrieval-augmented generation (RAG), augments prompts with custom data — typically from your internal data or knowledge bases. Some GenAI technologies, such as Microsoft Copilot, provide out-of-the-box RAG capabilities that help extend the model with internal data. HR leaders should work with their IT

counterparts to evaluate any RAG-prompting claims by AI and HR technology vendors to ensure confidential and personal data is protected.

Adopting this approach without these built-in capabilities entails significant additional costs and demands complex alterations to the technical architecture. The knowledge and expertise required for implementing the necessary IT setup for such advanced approaches are quite rudimentary within most enterprises. However, some organizations seeking to differentiate by building out enterprisewide employee-facing services may take on this effort and include select HR domains in a broader scope.

HR leaders should also note that extending via data retrieval is a common deployment approach for HR technology vendors who embed GenAI capability in their solutions. HR leaders should include questions about GenAI partners when speaking with their vendors to evaluate risks and understand dependencies.



Extend GenAI Models via Fine-Tuning

The “extend via fine-tuning” approach allows you to customize a pretrained GenAI foundation model to focus on a narrower subject or specific goal. For example, a fine-tuned model can excel in HR-related tasks that incorporate specific corporate policies. Fine-tuning trains the model using a curated dataset of sample inputs and desired outputs while adjusting the weights of the parameters that influence how the model processes those inputs and formulates outputs. A large language model (LLM) often uses millions of parameters, so this approach should not be taken lightly. Like extending via data retrieval, this approach comes with significant costs and complexities and requires staff who are skilled in data science and AI. Additionally, as new HR tasks or requirements emerge, ongoing maintenance and additional fine-tuning may be necessary to ensure the model continues to perform effectively, adding to the complexity and cost of scaling.

Some HR technology providers may provide this fine-tuning themselves. HR leaders should work with their IT counterparts to evaluate any claims of fine-tuning by HR technology vendors to ensure personal data is protected.

Build Custom Foundation Models

The “build” approach entails creating custom foundation models from scratch, fully tailored to the organization’s data and business domains. This approach provides the highest level of control and alignment with specific use cases. HR functions are highly unlikely to adopt this approach, however, due to the substantial costs and complexities involved, as well as the rapid pace of innovation in the HR technology market.

Considerations for GenAI Deployment

Most HR teams will prefer the consume and embed approaches to GenAI deployment, as these will enable them to leverage prebuilt GenAI solutions. Understanding the other approaches can help HR leaders become better consumers of vendor-provided solutions. However, certain situations may necessitate a deeper integration of GenAI with an organization’s data to achieve domain-specific results. In such cases, the extend via data retrieval approach can align more closely with specialized HR requirements, especially if your GenAI technology vendor

provides built-in RAG capabilities. HR leaders should work with IT to understand if it aims to invest in extend or build approaches and can incorporate HR use cases. In exceptional cases, HR may have an embedded AI and application design team within the HR function that can tackle these efforts themselves.

Prepare for GenAI Deployment

To prepare for deploying GenAI capabilities in HR, every organization needs to do two things. First, engage with your existing HR tech platform vendors to gain insights into their product roadmap regarding GenAI opportunities. Next, collaborate closely with your technology teams to assess these vendors' readiness, considering both technical capabilities and their alignment with your organization's specific requirements, compliance and

security framework. Don't hesitate to ask challenging questions to vendors and seek concrete examples of how their tools perform in HR contexts.

Examples of questions to ask technology providers and IT teams include:

1. How is generative AI integrated into the solution, and what specific benefits does it provide beyond existing functionality?
2. What mechanisms will ensure the accuracy of the generated output, and how can we actively monitor and verify its correctness?
3. How are prompts defined, and what control do we have over the input for these prompts? For instance, when generating responses to common HR inquiries, what details about an employee are incorporated into the prompt to ensure a tailored and contextually relevant answer? Similarly, when creating marketing content, what specific data about a candidate or our employer brand guidelines is included in the prompt?
4. What version and specific features of your solution must we use before deploying generative AI?
5. How does the tool access and interact with our HR data and knowledge bases? What security measures safeguard this sensitive information?
6. How can we jointly engage in corrective actions if a generative AI tool provides false information?
7. How does the system manage highly sensitive HR data? Are certain use cases restricted to ensure data protection? How do we prevent misuse, such as including overly sensitive information in prompts?

Work closely with your IT department to understand their investments in GenAI tools. This alignment with existing IT assets can uncover synergies and maximize the value of GenAI within your HR operations. However, while exploring GenAI solutions, exercise caution to avoid overpromises by vendors. You must clearly understand the capabilities and limitations of the GenAI tools you're considering.

Note 1: Key Terms

Generative AI is the ability of a computer program to create brand-new, completely original variations of content, including images, video, music, speech and text. It can also improve or alter existing content, create new data elements and create novel models of real-world objects, such as buildings, parts, drugs and materials.

Foundation models (mostly large language models) are trained on a broad set of unlabeled data that can be used for different tasks with additional fine-tuning. They are called foundation models because of their critical importance and applicability to a variety of downstream use cases due to large-scale pretraining of the models. Foundation models are trained on general data sources in a self-supervised manner, which can then be adapted to solve new problems. Foundation models are based mainly on transformer architectures, which embody a type of deep neural network architecture that computes a numerical representation of text in the context of surrounding words, emphasizing sequences of words. Although they are primarily based on transformer architectures today, this may not always be the case in the future.

Prompt Engineering is the discipline of providing inputs, in the form of text or images, to generative AI models to specify and confine the set of responses the model can produce. The inputs prompt a set that produces a desired outcome without updating the actual weights of the model (as done with fine-tuning). Prompt engineering is also called "in-context learning," where examples are provided to guide the model further.

¹ Benchmark With Gartner: Special Edition — Navigating the Workforce Impact of Generative AI (21 June 2023). This webinar was part of a monthly series of live benchmarking sessions covering the top internal and external concerns that impact key talent decisions HR leaders need to make. These sessions enable HR leaders to discuss stands their peers are taking on timely, critical topics. Attendees represented a spectrum of industries globally with a focus on North America. Results from these sessions were compiled into a final benchmarking report.

Connect With Us

Get actionable, objective insight that drives smarter decisions and stronger performance on your mission-critical priorities. Contact us to become a client:

U.S.: 1 855 811 7593

International: +44 (0) 3330 607 044

[Become a Client](#)

Learn more about Gartner for HR

gartner.com/en/hr

Stay connected to the latest insights



Attend a Gartner conference

[View Conferences](#)

© 2024 Gartner, Inc. and/or its affiliates. All rights reserved. CM_GBS_2867240

Gartner[®]