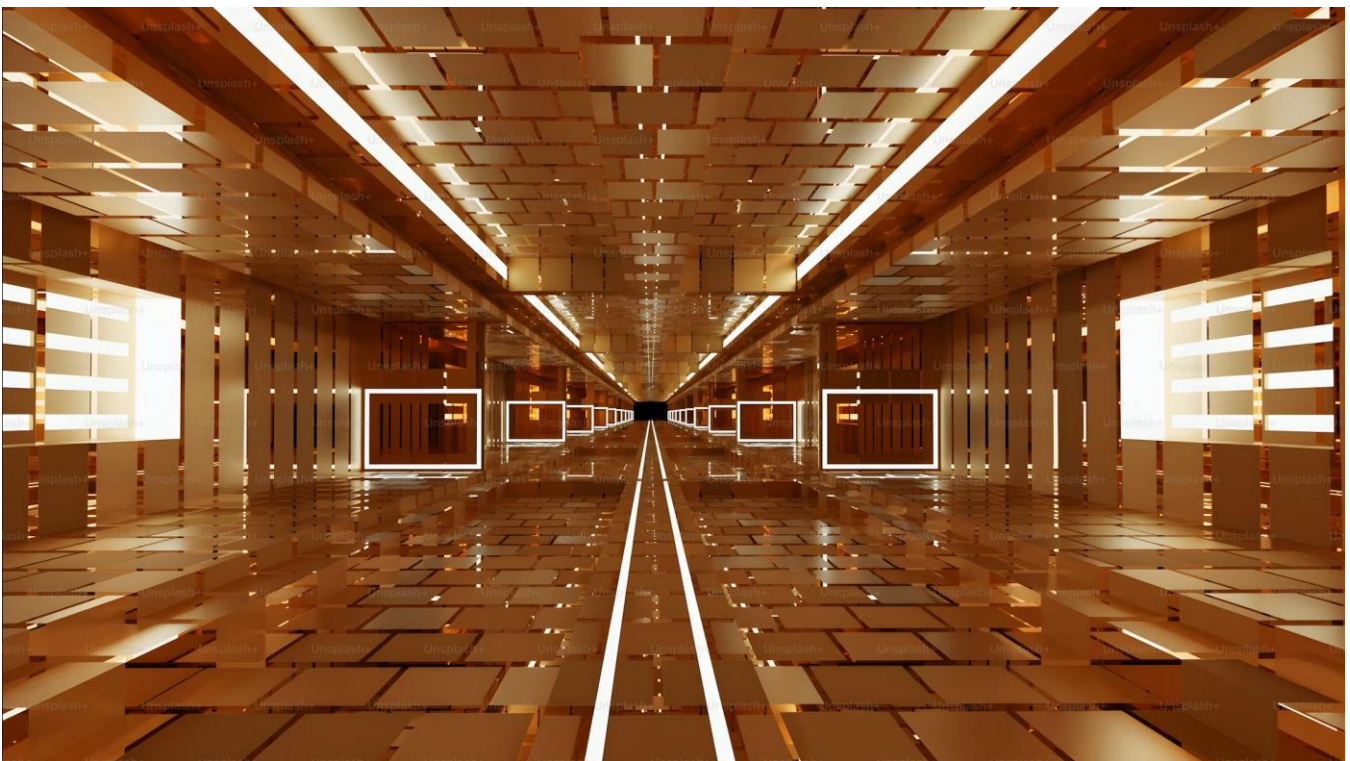


The HR One-Stop Shop Is Finally Within Reach

But only with a specialist-first, governed agent architecture

André Fortange & Volker Jacobs, Insight222 · 04 May 2026



This article offers CHROs and senior HR transformation leaders a practical decision frame for building a credible HR one-stop shop with governed Lead and Expert Agents.

It draws on TI People's HR Transformation Leaders Program, consulting practice, targeted practitioner interviews, and a structured evidence review.

IN THIS ARTICLE

- What makes an AI-enabled HR service model credible
- Why the debate has changed
- Three architectures are now credible
- Specialist-first is the only credible path
- What this means for HR's operating model
- The platform debate: the right stakes, wrong question
- Governance belongs in the design, not after it
- Where the economics actually sit

Most HR leaders do not start with an architecture question. They start with a service problem: employees cannot find the right answer, managers do not know where to send a request, and HR teams spend too much time moving work between systems, policies, tickets, and experts. The promise of AI in HR service delivery is not simply a better chatbot. It is a simpler way for employees to ask for help and a more reliable way for HR to route, resolve, and govern that work behind the scenes. This is the problem a Lead Agent architecture is meant to solve: one clear entry point for the employee, connected to specialist services that know the process, the policy, the data, and the limits of automation.

In plain terms

In this article, “front door” means the single place where employees and managers go first with an HR question or request. A “Lead Agent” is the AI layer that understands the request, asks clarifying questions if needed, and routes the work to the right specialist service, workflow, system, or human expert. “Expert Agents” are narrower capabilities behind the scenes, for example for leave, payroll, benefits, learning, or employee data that know the relevant rules, access the right systems, and operate within clear boundaries.

Summary

For years, HR leaders have wanted HR service to feel simple: one place where employees can ask a question, start a process, and get a reliable answer or resolved outcome without knowing which system, policy, or HR team sits behind it. The ambition was never the problem. The difficulty was making that experience work reliably across fragmented HR systems, local rules, sensitive data, and human exceptions. Across our consulting work and targeted interviews with HR technology and HR transformation experts in the field, a clearer picture is now emerging: a Lead Agent model is credible for HR, but only if it is built as a front-door routing layer above narrow specialist services, under existing user permissions, with auditable outputs and clear human handoffs. The bigger implication is not just better service technology. It is a redesign of how HR work is organised.

Key Takeaways

- A **Lead Agent is viable when it acts as the first point of contact and routing layer** — not as an all-knowing HR super-agent.
- The **most defensible implementation path is specialist-first**: prove quality and trust by solving one or two high-value services first, then add a broader front door.
- The **decisive architecture choices sit below the interface**: source-system access, permissions, auditability, and human handoff logic.
- The **platform question matters, but less than most vendors suggest**. Judge platforms by the capabilities they contribute, not the front-door role story they claim.
- The **larger implication is operating-model redesign**: less first-line transaction work, more product ownership, knowledge governance, analytics, and exception handling.



What makes an AI-enabled HR service model credible

Distinctive value of this research is the degree of convergence across practitioners with different system landscapes, different vendor positions, and different starting points.

A single starting point matters. Employees and managers should not need to know whether their issue belongs in SuccessFactors, Workday, ServiceNow, SharePoint, payroll, or a local HR team. They should be able to describe the problem once and be guided to the right answer or resolution path.

The front door should route, not pretend to know everything. The Lead Agent interprets intent, disambiguates, and hands work to the best specialist service, workflow, or human path. That is its job. It should not attempt to answer everything directly.

Specialist depth beats generalist breadth in HR. The safest path is narrow domain logic, live system access, and clear boundaries. HR has too many high-stakes edge cases for a shallow generalist model to be trusted.

Document retrieval is not service resolution. Finding and summarising a leave policy is information access. Reporting an employee's current leave balance is already a service, because it depends on live system data. True service resolution goes further: confirming eligibility, securing manager approval, booking the leave in the source system, and notifying the employee once it is complete.

Governance is a design question, not a policy wrapper. Permissions, answer quality, and traceability are not late-stage controls. They are part of the architecture itself.

The core must be clean enough to connect. Agents expose the quality of the systems behind them. Where data is incomplete, process ownership is unclear, or exceptions are undocumented, agent failure is predictable — and faster to distribute than manual failure.

"Output quality is our biggest risk because the stakes in HR are fundamentally different from IT. A wrong answer about access rights is an inconvenience; a wrong answer about someone's retirement entitlement could lead them to make life-altering decisions. The quality bar for HR is therefore much higher."

SVP HR Technology and Services, Global Technology company

"You only need the general layer once you have at least two specialist cases to orchestrate. Before that, it adds little value."

Nikolaus Krasser, CEO, Pentos AG

Why the debate has changed

What has changed is not that HR suddenly needs a different ambition. It is that the underlying technology has crossed a practical threshold. Earlier digital HR tools could answer isolated questions. The new generation can receive a request, coordinate across multiple steps and systems, and return a coherent outcome. Three developments have shifted what is actually buildable.

AI that completes tasks, not just answers questions. Earlier AI tools responded to one question with one answer. The new generation can receive a request, determine that resolving it requires multiple steps across multiple systems, execute those steps in sequence and return a completed outcome. For HR, this is the difference between a system that tells an employee how many days of leave they are entitled to and one that tells them how many they have left, what is already booked, and confirms the new booking in the source system. That difference is not a better interface. It is a different category of capability.

Ethan Mollick describes the same shift as a move from co-intelligence to managing AIs: the user is no longer prompting for an answer, but directing an agent through a task that may span systems, steps, and time. His work on interfaces reinforces the same point: value depends not only on model capability, but on whether the agent fits naturally into how people already work.

Platforms that can now talk to each other. The cost and complexity of connecting an HR service layer to the platforms and systems behind it has fallen significantly. New standardised connection technology (specifically Model Context Protocol (MCP) and Agent-to-Agent (A2A) protocols, both introduced in 2024–2025) means a Lead Agent can reach across SAP, Workday, ServiceNow, and others without rebuilding each connection from scratch. Organisations no longer have to choose between their existing platforms. The connections are becoming more standardised; the architecture design is not.

Quality that is now reliable enough in scoped service areas. IBM's AskHR system has sustained a rate of 94% of queries handled without human involvement across 2.1 million conversations per year, spanning Workday, SAP SuccessFactors, and Concur. In carefully scoped service areas, with strict guardrails in place, the quality bar is now reachable. This does not mean a general-purpose HR brain has arrived. It means that reliable, employee-facing service delivery is achievable — for those who build the foundations to support it.

These developments do not mean a single “HR brain” has arrived. They mean a more modest but more useful architecture is now plausible: one employee-facing routing layer, many narrow specialist services behind it, and source systems that remain the operational truth.

"We built together with ServiceNow the connection between Now Assist and Joule, so both systems communicate consistently and data from SAP and SuccessFactors is available system-wide."

Head of HR IT Strategy & Solutions, Large Retail Conglomerate

Three architectures are now credible

For a CHRO, the architecture question is not primarily technical. It is a choice about how simple HR should feel to employees, how much control HR wants to keep over service quality and governance, and how dependent the function is willing to become on any one platform vendor.

HR leaders now face three credible architecture patterns. They should be treated as design options for different contexts not as competing ideologies and evaluated against the organisation's specific system landscape, governance requirements, and service ambitions.

	Architecture 1	Architecture 2	Architecture 3
	Embedded AI	Control-Tower	Lead + Expert Agents
Best fit	One-vendor or homogeneous HR stacks	Companies with a strong existing service layer	Heterogeneous landscapes needing cross-system composability
Main risk	Vendor lock-in; weak cross-platform reach	Platform dependency; new bottleneck risk	Architecture discipline; governance design required
Vendor examples	SAP Joule / SuccessFactors, Workday Illuminate / Workday Agent System of Record	ServiceNow AI Control Tower + Agent Fabric	IBM AskHR / watsonx Orchestrate; builds enabled by MCP and A2A connections
What the CHRO decides	Whether to deepen the current platform vendor relationship or preserve flexibility across systems	Whether to invest in a service layer that governs AI coordination across HR	Whether to build orchestration as an organisational capability rather than purchasing it from a platform vendor

Exhibit 1: Three HR agent architecture patterns

Architecture 1 (Embedded AI) is frictionless within its own platform but structurally limited at the platform boundary. For heterogeneous landscapes, which describes most large European employers, it is insufficient on its own. Architecture 2 (Control-Tower) offers centralised governance and cleaner observability, but carries the risk of creating a new form of platform dependency. Architecture 3 (Lead + Expert Agents) is the most composable option: orchestration becomes an organisational capability rather than a platform feature.

The practical implication is that Architecture 3 is not a rejection of Architecture 2. It is a more modular variant. Organisations can implement elements of both simultaneously.

Specialist-first is the only credible path

The most important correction is also the simplest: the front door should not try to solve everything itself. Its value lies in receiving the request, clarifying intent, and directing work to the right specialist capability. If there is only one specialist service, there is no reason to build a broad routing layer yet. If there are no trustworthy specialist services behind the front door, the front door becomes a confident but shallow generalist, which is worse than nothing in HR. Architecture 3 in particular demands this discipline: composability only creates value when the components being composed are genuinely capable.

The real architecture sits below the chat window. The visible interface is not the hard part. The hard part is: which domain is narrowly scoped, which source system is authoritative, what live data is required, how permissions are applied, and where a human must take over.

"The employee mainly doesn't want to know how many vacation days they're entitled to on an annual level. They want to know how many days are left, how many are planned, and do they have days left which they can book. That requires a real-time API, not a policy document."

Sebastian Zimmermann, HR Data & Knowledge Governance Manager, Lonza

The most important distinction is between document retrieval and service resolution. A system that can find a leave policy is useful. A system that can tell an employee how many days they have left, what is already planned, and whether they can book now has crossed into real service resolution. That requires live access to source systems, not just a knowledge repository.

Five Expert Agent design principles carry consistent support across the evidence, reordered here to lead with what matters most to employees and to HR's accountability.

- **Build explicit handoff logic before you build the automation.** Define in advance which cases stay automated, which require human augmentation, and which must route immediately to a person. In HR, ambiguous or high-stakes requests like pension queries, disciplinary matters, emotionally sensitive conversations,

escalate too fast for a shallow boundary to hold. The handoff is not a fallback. It is a designed feature of a trustworthy service.

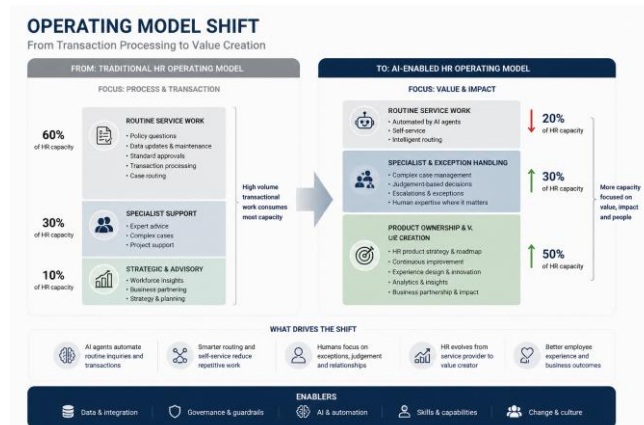
- **Treat answer quality as a trust issue, not a product quality issue.** Wrong answers about parental leave, partial retirement, or pay entitlement cause real harm. An employee acting on incorrect advice may face material personal and financial consequences. Quality thresholds must precede production rollout and must be defined before build begins.
- **Keep scope narrow and domain-specific.** Domain depth improves quality and auditability. A specialist agent that knows one service area well is more reliable and more governable than a generalist agent that knows many areas superficially.
- **Connect via real-time access to source systems, not static documents.** The difference between retrieving a policy document and confirming a live leave balance is a live system connection. Without it, service resolution is not genuinely possible, only information access.
- **Act under the requesting user's own permissions.** This single design choice resolves GDPR data minimisation, access control, and audit simultaneously. The employee sees what they are already permitted to see. Nothing more needs to be built.

Sequencing from specialist depth before general orchestration maps directly to the phased approach articulated in the HBR and Google Cloud blueprint for enterprise-wide agentic AI transformation: establish reliable specialist capabilities first, then build the orchestration layer that connects them.

A further discipline, designing against agent sprawl so that point solutions can be coherently orchestrated as the architecture scales, sits at the portfolio level, but is equally non-negotiable once more than one or two Expert Agents are in play.

What this means for HR's operating model

Architecture decisions cannot be separated from operating-model implications. The likely outcome of a mature agentic HR architecture is not the disappearance of HR, but a redistribution of effort and a materially different set of role demands.



First-line work will shrink as routine queries move into agent-

handled self-service. IBM AskHR's benchmark, a 75% reduction in support tickets and a 94% query containment rate across 2.1 million conversations per year, provides a reference point for what that shift looks like at scale. Complex, emotionally sensitive, or legally ambiguous cases will remain human. The function does not disappear; it moves upward in complexity while narrowing in volume.

The pressure on the classic HR business partner model is more structural. Where HRBP roles are already fragmented and operational, agentic service delivery is likely to expose that weakness rather than create it. The stronger pattern emerging from practitioner evidence is a sharper split between routine service support, deep expertise, and product-style ownership of employee problems.

BCG's AI at Work research points in the same direction: among heavy AI adopters, 29% expect fewer traditional entry-level roles and 43% expect greater demand for generalists who can manage human-agent teams. These figures should not be read as precise forecasts for every HR function, but they reinforce the same operating-model logic: less routine handling, more orchestration, judgment, and cross-boundary problem solving.

The customer success model deserves its own framing. In place of a senior generalist embedded with each business unit, the emerging pattern points to a two-tier structure: more scalable customer-success resources handling repeatable coordination and relationship management, supported by a bench of real subject-matter experts available for escalation. This is not a reduction in HR's ambition - it is a more honest description of where expertise should sit and how it should be

deployed. For many organisations, it implies a significant redesign of what the HRBP role actually does.

The sequencing question matters: organisations do not need to redesign HRBP roles before deploying the first specialist agent. The redesign becomes necessary and much easier to justify internally once the first services are live and the volume of routine queries visibly shifts. Build first; let the evidence make the case for role redesign.

"We expect many first-level support roles to disappear. The work will shift toward subject-matter expertise, KPI interpretation, training, and helping leaders understand what to do with the data."

Head of HR Technology and Services, Global Software company

HRIT faces an equivalent challenge: a deployment mindset is no longer sufficient. Organising around employee and manager problems, rather than system features or vendor roadmaps, is the more demanding and more necessary task.

A dimension that receives less attention than it deserves: HR's responsibility does not stop at its own function. As agentic architectures mature within HR, the function is increasingly well-placed to share what it has learned about governance, human handoff design, and responsible deployment with other functions facing the same choices. The CHRO and HR leadership have a dual role: deploying AI within HR and shaping how AI is deployed responsibly across the enterprise. That broader stewardship mandate, informed by McKinsey's research on HR's transformative role in an agentic future, is becoming an explicit part of the CHRO's agenda, not an afterthought.

Current emphasis	Likely shift	Growing roles / capabilities
High-volume first-line handling	More self-service and automated containment	Quality assurance, exception handling, analytics
Generalist HRBP coordination	Sharper split between routine support and deep expertise	Customer success model, strategic specialist roles
Siloed process ownership	Product-style ownership of employee problems end to end	HR product manager, service designer
Distributed, unmanaged knowledge	Governed knowledge as a strategic asset	Knowledge owner, content steward
Tool deployment mindset (HRIT)	Continuous service improvement; product thinking replaces deployment logic	AI solutions architect, platform product manager

Exhibit 2: Operating model shifts in an agentic HR architecture

TI People's research quantifies the shift clearly: in a 30,000-employee company, AI can free up 29% of HR capacity, equal to 66.4 FTEs. The question is not whether that capacity is freed, it is where it goes. Operational roles offer the biggest automation upside, while strategic roles face the greatest redesign pressure. The research identifies four emerging roles: HR Product Manager, HR AI Ethicist, HR Service Experience Designer, and HR Data Scientist. This points to a more product-oriented operating model as the structural response to AI pressure from both ends. The organisations that manage this transition well will not simply do less with more. They will direct freed capacity into the higher-value work that agents cannot do.

Two new structural roles reflect enduring requirements in an HR function that depends on agentic service delivery. The knowledge owner is responsible for the accuracy of the knowledge base specialist agents rely on. The agent designer translates service problems into workflows and usable capabilities. Neither role is optional at scale.

The platform debate: the right stakes, wrong question

The debate over which platform should sit at the centre of HR's agent architecture is real. But it is frequently framed in a way

that does not serve HR leaders well. The question CHROs are most often asked is: which vendor wins? That is the wrong

question. The right question is: which service problems does your organisation need to solve, and which platform capabilities are genuinely necessary to solve them?

ServiceNow, Workday, and Microsoft each position themselves as the natural front door for work. Each has a commercial interest in becoming the orchestration layer through which other capabilities are routed. The more useful executive question is not which platform wins. It is where the key capabilities should sit: knowledge management, case management, workflow and approvals routing, access logic, and the employee-facing interface. Some of these are genuinely essential in any architecture. Others are important but replaceable. None of them automatically belong to one vendor.

"Workday sees itself as the place to start. So does ServiceNow. So does Microsoft. How willing are they to hand over the lead role? Will there be one central solution as a one platform to go? Which one will it be? That's a political question as much as a technical one."

Head of HR Data Governance, Global Industrial Manufacturing

The practical rule is straightforward: judge every platform by the capabilities it contributes, specifically knowledge, ticketing, workflow, approvals, identity, observability, and not by the front-door role the vendor wants it to own. The right architecture is the one that preserves governance, avoids unnecessary duplication, and creates the simplest coherent experience for the employee.

This is also where Jason Averbook's composable-architecture argument and Mercer's agent system-of-record framing are useful. Organisations need a clear management and governance layer for agents, but not necessarily one monolithic service platform that owns every function and every interaction.

Microsoft's position deserves specific mention. With Copilot embedded across Microsoft 365 and Teams (where many employees already spend most of their working day) Microsoft carries structural advantages in the employee-facing layer that are independent of HR's own platform choices. Whether that proximity translates into genuine HR service capability depends on the depth of HR system integration behind it, which remains uneven across organisations. For HR leaders with a predominantly Microsoft environment, this is a variable worth assessing explicitly rather than assumed.

Governance belongs in the design, not after it

Governance is not something to add after the architecture is built. In HR, governance is part of the design. Three principles matter most.

Let your existing permissions do the governance work. The safest architecture is one where the AI acts on behalf of the employee making the request, using only what that employee

is already permitted to see. This resolves GDPR data minimisation, access control, and auditability in a single design decision. Hence no new permissions model is required, no centralised data copy is needed. The risk of the alternative is significant: when HR data is consolidated into a new repository to make it more accessible to agents, every existing access structure is broken and must be rebuilt from scratch. At greater cost, greater fragility, and far greater GDPR exposure.

Treat answer quality as a compliance issue, not just a product quality issue. In payroll, pension, leave, and similar domains, the real risk is often not privacy. It is the harm caused by a plausible but wrong answer. An employee acting on incorrect advice about parental leave or partial retirement may face material personal and financial consequences. Under the EU AI Act, several employment-related AI use cases including hiring fall into high-risk categories, with obligations around documentation, risk management, human oversight, and conformity assessment. The precise implementation timeline may shift, but the direction of travel is clear: these requirements should be built into the architecture from day one, not retrofitted.

"We are seeing incorrect information come out even now due to incorrect meta search tags, incorrect translations and incorrect or outdated data. An employee acting on wrong advice about their pension or leave entitlement can cause serious personal harm. The quality bar for HR remains high and a critical element."

Sebastian Zimmermann, HR Data & Knowledge Governance Manager, Lonza

In co-determined environments, works council framing must come before scale deployment. This is a sequencing requirement, not a communications exercise. A Framework Works Agreement (Rahmenbetriebsvereinbarung) must be in place before broad rollout. Under §87(1) No. 6 of the German Works Constitution Act, any IT system capable of monitoring employee behaviour requires co-determination regardless of intent. Deployment without agreement can be blocked by court order.

Bitkom's guidance on AI and co-determination and Orrick's employment-law analysis point to the same practical conclusion: early agreement is an architectural sequencing requirement, not late-stage stakeholder management.

"If the agent goes into the system with the permissions of the person making the request, very little can go wrong. The moment you centralise data into a lake, you break all existing access structures and may need to rebuild them from scratch."

Nikolaus Krasser, CEO, Pentos AG

Where the economics actually sit

The economics are frequently misunderstood. **The primary cost is not the Lead Agent itself. It is the invisible work behind it:** integration, knowledge cleanup, testing, exception

design, and adoption. BCG's framing is direct: 10% of agentic transformation cost sits in the technology layer, 20% in data, and 70% in people and process redesign. Organisations that budget for the 10% and discover the 70% mid-programme generate exactly the failure rates Deloitte documents: over 40% of agentic AI projects risk cancellation by 2027 due to unanticipated cost and complexity.

IBM IBV makes a parallel point: weak foundations, poor interoperability, and insufficient model-management discipline are often the sources of agentic AI failure, not weak models alone. That is why the economic issue is less the price of the front-end agent than the readiness of the systems, data, permissions, and processes behind it.

One frequently overlooked risk in framing the business case: if cost savings are presented primarily as headcount reductions, Finance is likely to capture the benefit before HR can reinvest it in higher-value work. The more resilient framing treats agentic service delivery as capacity redeployment freeing HR from routine handling so it can focus on the work that only people can do. That framing is both more accurate and more defensible in the boardroom.

A practical value logic helps calibrate expectations. A narrow HR service becomes attractive when four conditions align: high repeat volume, bounded judgment, expensive human handling today, and source-system access that is realistically achievable. Consider a service generating 30,000 requests a year. If one-third can be safely handled or materially shortened, and each case saves eight minutes of handling time, the organisation removes roughly 1,300 hours of manual work, before counting quality gains, faster cycle times, or improved employee experience. The business case appears quickly in high-volume services, but only if integration and quality conditions are already in reach.

Case: IBM AskHR

IBM AskHR spans Workday, SAP SuccessFactors, and Concur. IBM started with domain-specific specialists, not a general Lead Agent. The published cost reduction took four years. A useful calibration against faster-build narratives.

94%	2.1M	40%
queries handled without human involvement	conversations per year	cost reduction over 4 years

The best public benchmark still points in the same direction. IBM AskHR reached scale with specialist domains, not with a generic front door first. And it took years, not weeks, to reach the published cost reduction (IBM, 2025). Fast ROI therefore comes from narrow scope and ready foundations, not from broad platform ambition.

"If the sub-agents work directly in the relevant systems under the requesting user's permissions, relatively little can go wrong. If you centralize everything into a data lake, you break all existing permission structures and have to rebuild them."

SVP HR Technology & Analytics, Global Technology company

A further economic risk deserves explicit attention: technology obsolescence. In a market evolving as quickly as agentic AI, a two-year implementation programme can begin to look dated before it has generated its expected return. The most resilient architectures are modular and composable, with shorter commercial commitments and the ability to swap components as model capabilities improve.

Mollick's work on the jagged frontier of AI reinforces this planning risk: general-purpose models can improve quickly enough to erode the advantage of some custom specialist builds. The implication is not to avoid specialist architecture, but to avoid locking it into components that cannot be replaced as model capability changes.

Four KPIs to Track from Day One

- Queries handled without human involvement — the primary measure of automated service value
- Escalation rate and reason — what kinds of cases are still routing to people, and why
- Mean resolution time — from employee request to confirmed outcome, end to end
- Answer accuracy rate — spot-check verification against source system data, on a defined cadence

Leadership Agenda

Leadership Agenda for 2026

The following steps reflect convergent evidence from practitioner interviews, external research, and consulting practice. They are sequenced deliberately: foundations and governance come before broad deployment.

- **Start with service problems, not platforms.** Use-case selection is the first strategic decision. The architecture should emerge from service problems, not from the desire to install a generic front door. Payroll queries, leave, parental leave, and benefits are typical starting candidates: visible pain, bounded scope, measurable value.
- **Build specialist services end to end before adding a general front door.** A routing layer creates value only when there is real specialist capability to route to. Starting with the orchestration layer first produces architecture without service value.
- **Clean the core before you connect it.** Poor data, unclear ownership, and undocumented exceptions do not disappear when AI is added but they become failure modes. Data quality, process ownership, and edge cases are part of the first use case, not a separate workstream.
- **Let your existing permissions do the governance work.** The AI should act on behalf of the employee making the request, using only what that employee is already permitted to see. This resolves GDPR, access control, and auditability simultaneously.
- **Treat answer quality as a trust issue, not just a product quality issue.** Wrong answers about parental leave, partial retirement, or pay entitlement cause real harm. Quality thresholds must be defined before build begins and must precede production rollout.
- **Judge platforms by capability contribution, not vendor narrative.** Knowledge management, ticketing, workflow, approvals, and identity all matter. None of them automatically earns the right to become the employee-facing front door.
- **Treat operating-model redesign as part of the build from day one.** Even a narrow first service changes role mix, handoff design, quality assurance, and knowledge ownership. Technology build and operating-model design must move together.
- **Build in a composable way and expect the landscape to change.** Model capabilities and vendor positions are evolving too quickly for rigid, long-range architectural bets. Shorter commercial commitments and the ability to swap components as capabilities improve are structural advantages, not signs of indecision.

For organisations in co-determined European environments: Works council framing and a Framework Works Agreement must come before scale deployment, not after. Under §87(1) No. 6 of the German Works Constitution Act, any IT system capable of monitoring employee behaviour requires co-determination regardless of intent. This is a sequencing requirement, not a communications exercise.

Conclusion

For HR leaders making these decisions in 2026, choosing which agentic service architecture to build is no longer a technical question: It is a leadership decision that shapes how HR creates value for the next several years.

The risk now is not moving too slowly or too quickly. It is building the wrong architecture by default, because the platform vendor told a compelling story and no one in HR set a different frame. That risk is avoidable. The architecture choices are now well enough understood and the practitioner evidence clear enough to make deliberately.

CHROs who engage with these decisions in 2026 will do more than choose a service model. They will set the terms on which their platforms, their teams, and their organisations approach AI. The function that masters the architecture question is better positioned to shape how AI is deployed responsibly across the enterprise, not just within HR.

The long-promised HR one-stop shop may finally be within reach. But reaching it requires more than a powerful front-end agent. It requires hard choices about architecture, governance, service quality, and operating model. Those choices belong in the CHRO's office, not in a vendor roadmap. The leaders who shape those choices in 2026 will determine how HR creates value for the next decade.

Sources

- Averbook, J. (2026). Now to Next. Substack / Mercer. January 2026.
- BCG (2025a). How Agents Are Accelerating the Next Wave of AI Value Creation. December 2025.
- BCG (2025b). AI at Work: Momentum Builds, but Gaps Remain. June 2025.
- Bitkom (2026). KI und Mitbestimmung: Betriebsvereinbarung für KI-Einsatz. February 2026.
- Deloitte (2025). Tech Trends 2026: Agentic AI Strategy. December 2025.
- European Union (2024). Artificial Intelligence Act, including high-risk AI obligations for employment-related use cases. July 2024.
- HBR and Google Cloud (2026). A Blueprint for Enterprise-Wide Agentic AI Transformation. February 2026.
- IBM (2025). Transforming HR support with agentic AI: IBM HR enhances employee experience with IBM watsonx Orchestrate.
- IBM IBV (2025). The Essential Guide to Agentic AI. November 2025.
- Jacobs, V. (2025). AI-Powered HR: How to Evolve the HR Operating Model for the AI Era. Insight222. March 2025.
- McKinsey (2025). HR's Transformative Role in an Agentic Future. November 2025.
- Mercer (2025). Heads Up HR: 2025 Is the Year of Agentic AI.
- Mollick, E. (2026). Claude Dispatch and the Power of Interfaces; The Shape of the Thing. Substack, March 2026.
- Orrick (2024). AI and German Co-Determination – What Employers Need to Know. September 2024.