

Assessing Delivery Judgment in Provider Teams

A Provider Method for Role-Based Competence

Phil Hatch | Akholi

October 23, 2026

Provider leaders should assess whether teams can manage the two-party agentic delivery model, not merely operate an AI tool. Competence includes recognizing authority limits, presenting decision-ready exceptions, preserving evidence, disclosing intervention, and coordinating correction with the buyer.

Define Competence Through the Delivery Decision

The provider leader should identify the judgments each role must exercise during live service. Operators, supervisors, commercial owners, specialists, and incident leaders need different capabilities. Generic prompt training does not establish that they can direct consequential work across a customer boundary.

McKinsey's organizational analysis emphasizes new development paths for judgment in human-agent work.¹ It is not a validated training curriculum. The provider should define competence from the actual service tasks, authority, evidence, and customer responsibilities.

For a hypothetical operator, the important skill may be recognizing that an agent's success message lacks evidence of completion, preserving the original case, and escalating with the information a buyer needs. Producing a better prompt is only one possible part of that response.

¹ McKinsey & Company, "AI Is Changing Work. Now It Has to Change the Organization," McKinsey & Company, October 5, 2026, "Operating model," "Talent," and "Enabling AI reinvention," <https://www.mckinsey.com/capabilities/people-and-organization/our-insights/ai-is-changing-work-now-it-has-to-change-the-organization>.

Assess Judgment Where Work Becomes Difficult

The study reports manual intervention, knowledge dependency, and delivery processes not adapted to agentic work.² It does not establish that every team lacks competence. It identifies consequential situations in which practical assessment can reveal what generic training misses.

Use ambiguous instructions, missing authority, failed tools, conflicting records, buyer delay, and a need for human correction. Ask the person to explain what they may do, what they must preserve, who decides, and what remains unresolved. Evaluate the action and evidence, not confidence or presentation alone.

Distinguish knowledge from operating conditions. A capable employee can still be unable to perform because the workflow lacks access, time, authority, or customer response. Assessment should identify organizational defects as well as individual development needs rather than blame staff for an impossible design.

Use a Delivery Judgment Assessment Record

Figure 1 proposes four capabilities: recognize, act within authority, communicate the decision need, and verify disposition. Define role-specific standards with the customer where relevant. Record observed evidence, limitations, feedback, and the work the person is authorized to perform.

Bain's technology-services analysis identifies changing talent and delivery models.³ The provider should connect those changes to practical competence. Training completion is an input; demonstrated ability to manage the service is the relevant operating result.

Include joint exercises with buyer roles for high-consequence work. A provider employee should know how to prepare an exception, but the buyer must also supply qualified authority. Test the handoff and response rather than evaluate each organization in isolation.

² Phil Hatch, *Agentic AI Outsourcing at Tier 1 Financial Institutions: Full Study Report*, version 1.8.2 (Akholi, forthcoming manuscript dated November 1, 2026), Part 5, "The Five Manual Alteration Methods" and "Disclosure of Manual Alteration," table 5.8, <https://doi.org/10.2139/ssrn.7164559>; Hatch, *Agentic AI Outsourcing*, Part 4, "Knowledge at Risk of Departure."

³ Megha Chawla et al., "Tech Services: Cracking the Code on AI-led Growth," Bain & Company, September 29, 2026, "At a Glance" and discussion of commercial models, <https://www.bain.com/insights/tech-services-cracking-the-code-on-ai-led-growth-technology-report-2026/>.

Figure 1. The delivery judgment assessment record



Author's proposed method; adapt to the service and applicable jurisdiction.

Source: Phil Hatch, proposed framework developed for this paper.

Maintain Competence as Scope and People Change

Refresh assessment after material changes to authority, tools, models, processes, or customer requirements. Staff moving between accounts should learn the new context and limits. A certification from one service should not automatically authorize a different consequential role.

EU, UK, Swiss, US, and other services may require specific qualifications, training, professional decisions, records, or employee protections. Qualified owners should determine relevant local conditions. Internal competence assessment should not be represented as external accreditation or legal authorization.

Protect truthful reporting. If staff are penalized for every disclosed intervention or uncertainty, the assessment can reward appearance rather than judgment. Evaluate appropriate escalation and evidence quality alongside accepted outcomes and avoidable error.

Authorize Roles on Demonstrated Capability

A reasonable objection is that practical assessment takes people away from delivery. Focus on the decisions that can materially affect quality, authority, or recovery, and integrate exercises into mobilization and change. The cost belongs in the service model where competence is necessary for delivery.

Before assigning a consequential role, require the person to demonstrate representative and difficult cases with the actual tools and escalation route. Provide feedback and supervised practice, then reassess gaps. Restrict unsupported authority until competence and operating prerequisites are established.

The provider should know who can manage which part of the two-party service and on what evidence. Delivery judgment is demonstrated through appropriate action under uncertainty. It is not established by attendance at an AI course or the ability to generate fluent output.

Evidence and Source Note

The study is a nonrepresentative cross-section of 850 vendor delivery-team respondents on 150 engagements serving Tier 1 financial institutions, collected during January–June 2026. Neither banks nor vendors corroborated the accounts. It establishes reported experience, not causation, legal breach, or market prevalence. Its application here is a mechanism to investigate, not an estimate of conditions in other industries. The author-supplied version 1.8.2 is dated November 1, 2026 and is cited as a forthcoming manuscript consulted October 23, 2026.⁴

This is an analytical decision paper based on the supplied study and selected primary public sources checked through October 23, 2026. The original contribution is a role-based delivery-judgment assessment record focused on cross-firm authority, evidence, escalation, and verified disposition. The proposed method has not been independently validated. Legal requirements depend on the entity, activity, jurisdiction, and effective provisions; supervisory observations, research, and professional commentary have distinct authority.

Bibliography

Chawla, Megha, Mark Brinda, Sushant Khandelwal, and Sandeep Nayak. “Tech Services: Cracking the Code on AI-led Growth.” Bain & Company. September 29, 2026. <https://www.bain.com/insights/tech-services-cracking-the-code-on-ai-led-growth-technology-report-2026/>.

Hatch, Phil. *Agentic AI Outsourcing at Tier 1 Financial Institutions: Full Study Report*. Version 1.8.2. Akholi. Forthcoming manuscript dated November 1, 2026. Author-supplied manuscript consulted October 23, 2026. <https://doi.org/10.2139/ssrn.7164559>.

McKinsey & Company. “AI Is Changing Work. Now It Has to Change the Organization.” McKinsey & Company. October 5, 2026. <https://www.mckinsey.com/capabilities/people-and-organization/our-insights/ai-is-changing-work-now-it-has-to-change-the-organization>.

⁴ Hatch, *Agentic AI Outsourcing*, Introduction, “Executive Summary” and “Evidentiary Basis and Its Limits.”