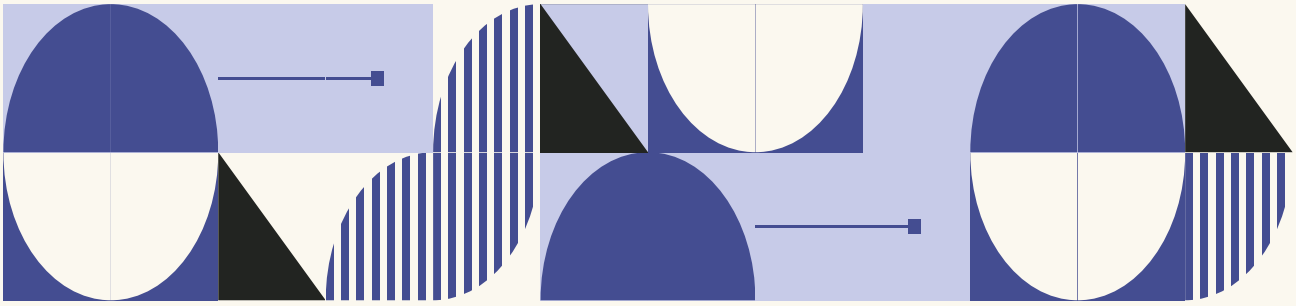
by Radion Barakov · © 2026

Matching Intelligence *to Work*



Who should do the work when AI can help, and when is a change worth it?

First illustrated edition

labourintelligencearbitrage.com

Contents

Orient	Opening note	03
	Executive summary	04
Diagnose	1. What changes when AI helps deliver a service?	05
	2. Defining a recurring type of work	07
Define	3. Routing work to people and AI	09
	4. When an arrangement meets the requirements	12
Apply	5. From assessment to decision	16
	6. Northstar: who should close supplier cases?	18
	7. Fableoak: sharing the work behind monthly reports	23
Challenge	8. Testing LIA: evidence and limits	25
	9. How LIA fits with existing approaches	27
Close	Closing note	29
Reference	Appendix A. Glossary	30
	Appendix B. L0-L5 diagnostic	33
	Appendix C. Assessment outline	35
	Appendix D. Evidence and source notes	37

Reading guide

For orientation, read the summary and the closing note.
For a delivery decision, start with Figure 1, then read the Northstar case and the assessment example.
For work shared across companies, see Fableoak.
For what a trial can establish, read Testing LIA.

© 2026 Radion Barakov · LIA

Opening note

As I watched AI find its way into service work, I kept coming back to a simple question. If we do the work differently, how do we know the result still does what people need it to do? A lot depends on what the organisation puts in place for the people doing the job. They need access to the right records, tools they're allowed to use and a clear sense of what they can decide. With those basics in place, someone can start by using AI for one part of the job and learn as they go. I was especially curious about services that draw on years of practice. AI might help pull the information together more quickly. It can still take experience to spot a missing record or see that two records tell different stories.

I'd look at what employees could do with AI and where an independent specialist or an experienced provider abroad could take on part of the work. Once those arrangements can meet the same standard, I'd want to know what each really costs. And if people spend less time gathering information, I'd want to think about where to put that time back. Some of it could go into giving difficult cases more attention. Some could give people room to practise with AI and get a better feel for when its output needs a closer look.

I wrote this paper to share how I think through those choices. I'd start with what the business needs from a service and how it's delivered today. Then I'd ask which parts AI could help with or complete, who would do the rest and who would be responsible for the result. What would change in the service people receive and in what it costs to deliver? The businesses in the paper are fictional. Their stories show how you might work through those questions in your own business.



© 2026 Radion Barakov · LIA

Executive summary

AI can find an invoice in seconds while the supplier still waits for an explanation. What would have to happen before the supplier agreed the case was finished?

Labour-Intelligence Arbitrage (LIA) starts with a question about work a business does regularly. Could we get the same result another way? It helps you work out what a good result looks like, who can make the decisions and what it really costs to get it.

AI can help people meet the standard or do the work with less effort. When options meet the same standard, the opportunity for **arbitrage** is the difference in what each usable result costs. A nearshore or offshore provider using the same platform may widen that advantage.

Key takeaways

- **Say what the service must deliver.** Checking an invoice, explaining why it's unpaid and deciding whether to pay it are three different jobs.
- **Separate the work from the choice of provider.** AI could gather the records while a person answers the question, whether that person works for you or for an outside provider.
- **Check what each option can handle.** Before switching, see whether the current service and the alternatives can meet the same reliability and timing requirements, including when records disagree.
- **Look beyond the fee.** A cheaper provider can cost your business more if your staff spend extra time checking answers or fixing mistakes.
- **Decide where saved time would help most.** Time freed up by AI can go into improving the service or giving people room to get better at using AI.

If an outside provider takes on the work, your own staff may still need to check answers and make decisions. To judge the change, look at what the supplier gets and what the business still has to do to deliver it.

Start with one piece of recurring work and ask what would make it worth doing differently. Two fictional cases show how to work through that question: Northstar compares providers, and Fableoak considers sharing work across companies. Figure 1 and the Northstar assessment example show how to start.

© 2026 Radion Barakov · LIA

1. What changes when AI helps deliver a service?

Whether a business sells products, provides or buys services, or organises the work behind them, AI raises a practical decision about who does the work and who manages it. Take a product company handling a warranty claim. AI might gather the purchase record and draft a reply. Someone still has to establish whether a replacement is owed and refer any decision they lack authority to make.

Statistics Netherlands (2026) found that in April about 36% of Dutch business-services firms were aiming for more automation to ease staff shortages. The survey included measures such as robots and AI support. In a separate March survey, the U.S. Census Bureau (2026) found that about a third of workers who had used AI at work in the previous week said it saved them one to two hours. For someone waiting on a warranty claim, what matters is how well it's handled and when they hear back. Time saved gathering records could give the handler more room to look into a difficult case or help the customer get a decision sooner.

Using AI also changes what the business needs to manage. Someone has to keep records available, check that results meet the agreed standard and deal with cases the arrangement cannot handle. Those responsibilities matter whether the work is done in-house, by a provider or through a combination of the two.

When AI changes the work, an hourly rate still tells the buyer what an hour costs. The full cost depends on how many paid hours, tools and checks it takes to deliver the result. If preparing a warranty case gets quicker but approval still takes the same time, the business may spend less effort on the case while the customer waits just as long. The customer may value a quicker answer, a reliable decision or better handling of a difficult claim. What the provider charges and what it costs to deliver are separate questions.

© 2026 Radion Barakov · LIA

Where AI helps and where it falls short

Brynjolfsson, Li and Raymond (2025) studied customer-support staff using an AI assistant at a software company. Less experienced and lower-skilled staff improved the speed and quality of their work. The most experienced and highest-skilled staff gained little speed, and the quality of their conversations fell slightly. The same tool can help people doing the same job in different ways.

In an experiment with Boston Consulting Group consultants, Dell'Acqua and colleagues (2026) found that GPT-4 helped on several consulting tasks. But when consultants had to combine interview notes and financial data to choose a brand for growth, those using GPT-4 were less likely to choose correctly. How much AI helps can change from one task to the next, even within the same kind of work.

Taken together, the studies suggest looking at both who will use AI and what they'll use it for. This paper starts there and asks a further question. If AI changes how part of a service gets done, what would make that change worthwhile?

What this paper examines

This paper examines how people, AI or a combination can deliver recurring work under the conditions a business actually has. It starts by defining what the work must deliver and who has authority to decide. It then looks at who or what could complete it, which options can meet the agreed standard, and how their full costs compare.

Two fictional cases show how that comparison informs a decision. Northstar considers who should handle supplier cases through closure, and Fableoak considers sharing report preparation across companies. The paper then looks at what evidence would show whether the method improves decisions and how it draws on existing approaches to work, cost and governance.

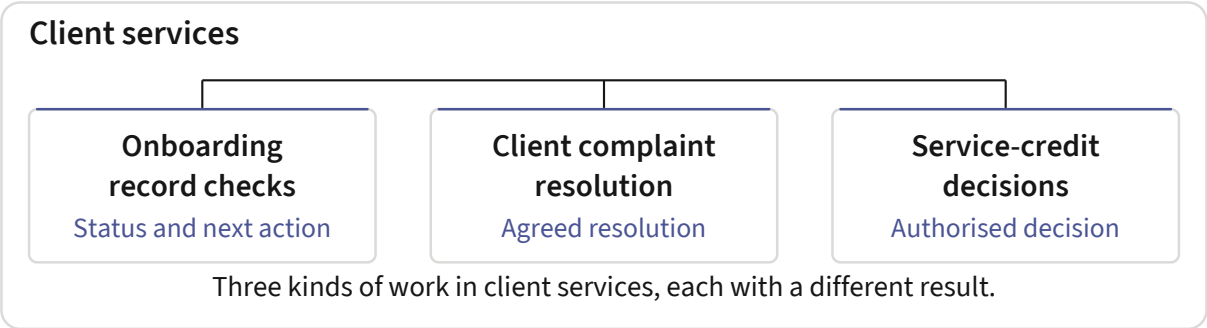
That's why it helps to separate jobs that sit together in a service. Checking an invoice, explaining why it's unpaid and deciding whether to pay it each produce a different result.

© 2026 Radion Barakov · LIA

2. Defining a recurring type of work

In client services, one person might check a new client's records, another might handle a complaint, and a manager might decide whether to grant a **service credit**. A service credit reduces what the client pays after a service problem. All three belong to the same **function**, the area of the business we call client services. Each leaves someone needing a different result.

Some jobs come up often enough to compare how they're done. LIA calls a recurring kind of job a **task class** when the cases and exceptions are similar, the work must deliver the same result, and the same rules apply about who can decide. Looking at the work this way helps show where it can get stuck and what still needs to happen. Revisit the class if recurring cases need a different result, different decision authority or materially different handling.



Three kinds of work in client services, each with a different result.

- **Checking onboarding records**
  A signed agreement can make the file look ready even when the delivery team is still waiting for a service schedule. The check should leave a record of what's missing and who will follow up, so the next person knows what still needs to happen.
- **Resolving client complaints**
  A quick reply can acknowledge a complaint while the client's problem remains open. What matters is whether the client's issue was resolved in line with the agreed service terms.
- **Deciding on an out-of-policy service credit**
  Someone can gather the facts and pass the request on, but the client is still waiting for a decision. The result they need is a decision from a manager with authority to approve or decline the credit.

© 2026 Radion Barakov · LIA

Four controls for how the work gets done

Quality threshold
Sets what the result must meet, including any deadline.

Exception budget
Limits extra cost from unplanned problems over a set period.

Escalation path
Names who takes cases needing more expertise or authority.

Review interval
Sets when to reconsider the arrangement.

When a service schedule is missing, the status record should show the gap by the agreed deadline. If the scope is unclear, the delivery lead takes the case. Any extra cost from repairing an unexpected error counts against the exception budget. At the monthly review, the team can look at how often the records were accurate and on time and how much extra work it had to absorb.

With the result and controls agreed, we can compare what people and AI could each do. Whoever does the check, the next person still needs the same reliable record.

© 2026 Radion Barakov · LIA

3. Routing work to people and AI

The onboarding check from Section 2 ends with a status record that shows what's missing and who will follow up. There's more than one way to get to that point. A coordinator could check the records by hand, use AI to prepare a draft or, if the workflow can meet the agreed standard, leave it to complete the record. LIA calls each way of getting the work done a **route**, including the support and checks it needs.

For routes involving people or AI, LIA uses three **completion destinations** to describe who or what delivers the result.

- **AI execution**
  An AI workflow completes the onboarding status record within agreed controls, with a named person responsible for the workflow.
- **Human-led delivery**
  AI can gather the history of a complaint, while the person dealing with it speaks to the client and resolves the issue.
- **Authorised human decision**
  An out-of-policy service credit is decided by a manager who has authority to approve or decline it.

A manager might also check a reply before it goes out. For the complaint, that check is one step towards resolving the issue. For the service credit, the manager's decision is what the client is waiting for.

Sourcing is choosing who provides the work. Employees, independent specialists and people working for a provider can do all or part of the same work under agreed conditions. If software that follows set rules can complete the result on its own, record it as a separate option and check it against the same requirements. One example is a check that flags an invoice mismatch. Appendix A brings these terms together.

© 2026 Radion Barakov · LIA

What the onboarding check needs to run

Suppose the coordinator uses AI to draft the status record and completes it after checking the source records. The tool needs access to the approved records and a current checklist. The coordinator needs to see the record behind each proposed status and know who can clear up a gap. With that setup, the tool can help find details and the coordinator can spend more time checking unclear items and arranging follow-up. The team can record corrections to see where the draft helps and where it creates extra work.

Table 1. The same onboarding check under different conditions. The situations are fictional. In each one, the status record shows what's arrived, what's missing or still unconfirmed, and who will follow up.

Situation	What's happening	What to find out
Manual check	Staff check records by hand.	Is the status record accurate and on time? What does the check cost today?
Missing records	Staff can't access some records.	Who can supply or verify them?
Unused AI tool	An approved tool is available.	Where, if anywhere, would it help?
AI draft under review	A coordinator checks the draft.	Can they verify it against the records?
Ready to try	Records, checks and roles are in place.	Does it meet the standard, and what does it cost?

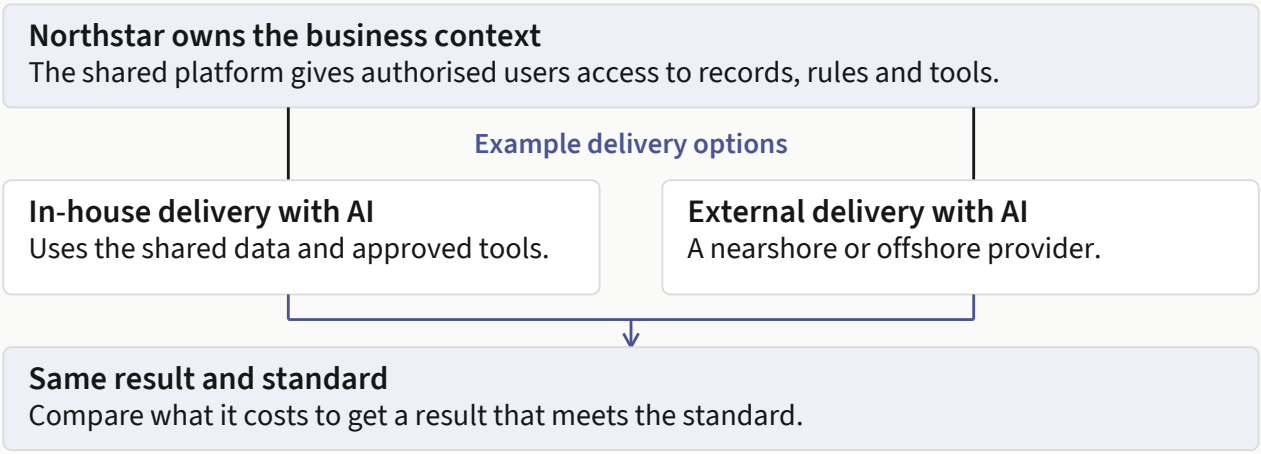
If an outside provider takes on the same check, a lower fee is one part of the comparison. The business may still spend time sharing records safely, answering questions, checking the results and fixing mistakes. The provider also needs access to the right records, support when something is unclear and room to make the decisions it's authorised to make. The arrangement has to meet the applicable working, contracting and data requirements.

The person running the assessment keeps the comparison record and discloses any financial interest in the recommendation. The named decision-maker can approve the change, keep the current arrangement, wait for evidence or authorise a limited test. The provider is responsible for the service it agrees to deliver.

© 2026 Radion Barakov · LIA

Different ways to deliver, one business context

Consider Northstar, a fictional business comparing ways to handle supplier cases. Its own staff and an outside provider would use the same records, tools and rules through a shared platform controlled by Northstar.



A qualified provider can build or maintain the platform. Staff and outside professionals can use the same data layer and approved tools, with access set for each role.

When records are incomplete or teams use terms differently, someone needs to map, clean and maintain them. That work has a cost. The records have to be reliable enough for the service.

What stays the same?

The decision flow shows what evidence each option needs. Once the options qualify, compare their full cost per accepted result.

© 2026 Radion Barakov · LIA

4. When an arrangement meets the requirements

A status record can look complete even when it marks a missing document as received. The layout can vary, but the next person needs an accurate account of what's missing and who will follow up, ready by the agreed deadline. Those requirements set the **quality threshold**. A result that meets them is an **accepted result**. Agree how reliably a route must deliver accepted results to qualify. Before assessing the options, set separate requirements if failures on certain cases or missed deadlines would have serious consequences. Some work has a deadline. Other work has to reach a checkpoint before the next step.

First check which options can meet the standard under the conditions they'll work in. Include the current service if it qualifies. Exclude an option that falls short and hold a proposed change when evidence is missing. Among options that qualify, start with the lowest **full cost per accepted result**. Keep total spending visible for the budget decision.

Sometimes the comparison brings out something worth paying more for, such as wider coverage. The decision-maker can make that benefit a requirement for every option or record why it's worth paying more for.

Comparing the full cost

Finance staff could check supplier invoices against purchase orders by hand, or use AI to draft the findings for someone to verify. Both routes have to deliver the same thing: a correct record of whether the invoice matches the order, with any discrepancy flagged by the agreed deadline.

Suppose the current route costs €960 for a batch of 100 invoices, including staff time, tools, review and corrections. Three checks are late, and one is on time but misses a price discrepancy. The business pays for the work on all 100 invoices and gets 96 checks that meet the standard.

What are we measuring?	Calculation
Cost per invoice received	€960 ÷ 100 = **€9.60**
Cost per correct, on-time check	€960 ÷ 96 = **€10.00**

The cost per correct, on-time check is **about 4% higher** than the cost per invoice received. The total bill is still €960. What changes is how much usable work the business gets for that money. That's the figure to compare across routes once each has shown it can meet the requirements, using the same period, workload, mix of cases and standard. If there are no accepted results, the route has no usable cost per accepted result.

© 2026 Radion Barakov · LIA

Four checks may seem like a small shortfall in one batch. If it happens month after month, or similar work has the same problem, the extra checking, corrections or delays can add up. That matters when deciding whether to keep the current route. Keeping it may avoid the cost and disruption of switching, while leaving the business with costs it already absorbs. A different route is worth considering if it can improve what the business gets enough to justify what it takes to change and run it.

What goes into full cost

Table 2. What goes into the cost of delivering a service.

Cost category	What it covers
People and service charges	Paid time or provider fees
Tools and AI support	Licences, usage and system connections
Coordination and control	Handoffs, supervision and decisions the business retains
Checking and correction	Review, unsuccessful attempts and rework
Setup and transition	Assessment, testing and moving the work
Keeping the service running	Staff support and updates to workflows, tools and connections

A provider’s fee may already cover tools, staff support or routine checks, while the business may still pay for secure access or spend staff time reviewing the work. Looking at what the fee covers helps the business count each cost once and see what still sits with it, including its share of any common costs.

A lower cost per result can leave spending unchanged. If a manager spends less time checking the work, payroll may stay the same even if some of that time becomes available for other work. It helps to keep track of how much time was freed up, how much the business could use elsewhere and whether spending changed. The business can then judge whether the difference is large enough to make a change worthwhile.

A team can put some of that time into changing how another part of the service runs. Someone who spends less time finding supplier records might use those hours to help the onboarding team check new cases. Record how the time was used, what still had to be paid for and whether the new check worked.

© 2026 Radion Barakov · LIA

Exceptions, escalation and review



Exception budget

Some cases take extra work even when the process is running as expected. If an invoice and a purchase order disagree, sending the discrepancy to someone who can look into it may be a normal part of the check. If an error means several checks have to be redone, that adds an unplanned cost. The **exception budget** sets how much extra cost the business is prepared to absorb over a given period, whether people or AI complete the work. The record can show what departed from the normal route, why, who authorised it and what it cost. That includes staff time to put it right or request more spending. The forecast allows for expected rework, and the review counts the actual cost once.



Escalation path

When a case needs to go to someone else, the reason matters. Someone who knows the records may be able to explain a discrepancy. A decision about whether to pay a disputed amount may need someone with authority to make it. The **escalation path** makes clear who takes over in each situation. If a legal, safety or authority breach requires an immediate stop or escalation, that rule applies regardless of how much budget is left.



Review interval

Checking individual results helps the business see whether the work is being done properly. The **review interval** gives it a point to step back and ask whether this is still a good way to do the work. A change in workload or a rise in correction costs might bring that review forward. If the business hasn’t agreed what counts as acceptable, what extra cost it can absorb, who takes unusual cases or when to review the work, those are questions to settle before making a change.

© 2026 Radion Barakov · LIA

Figure 1. Decision flow for task delivery

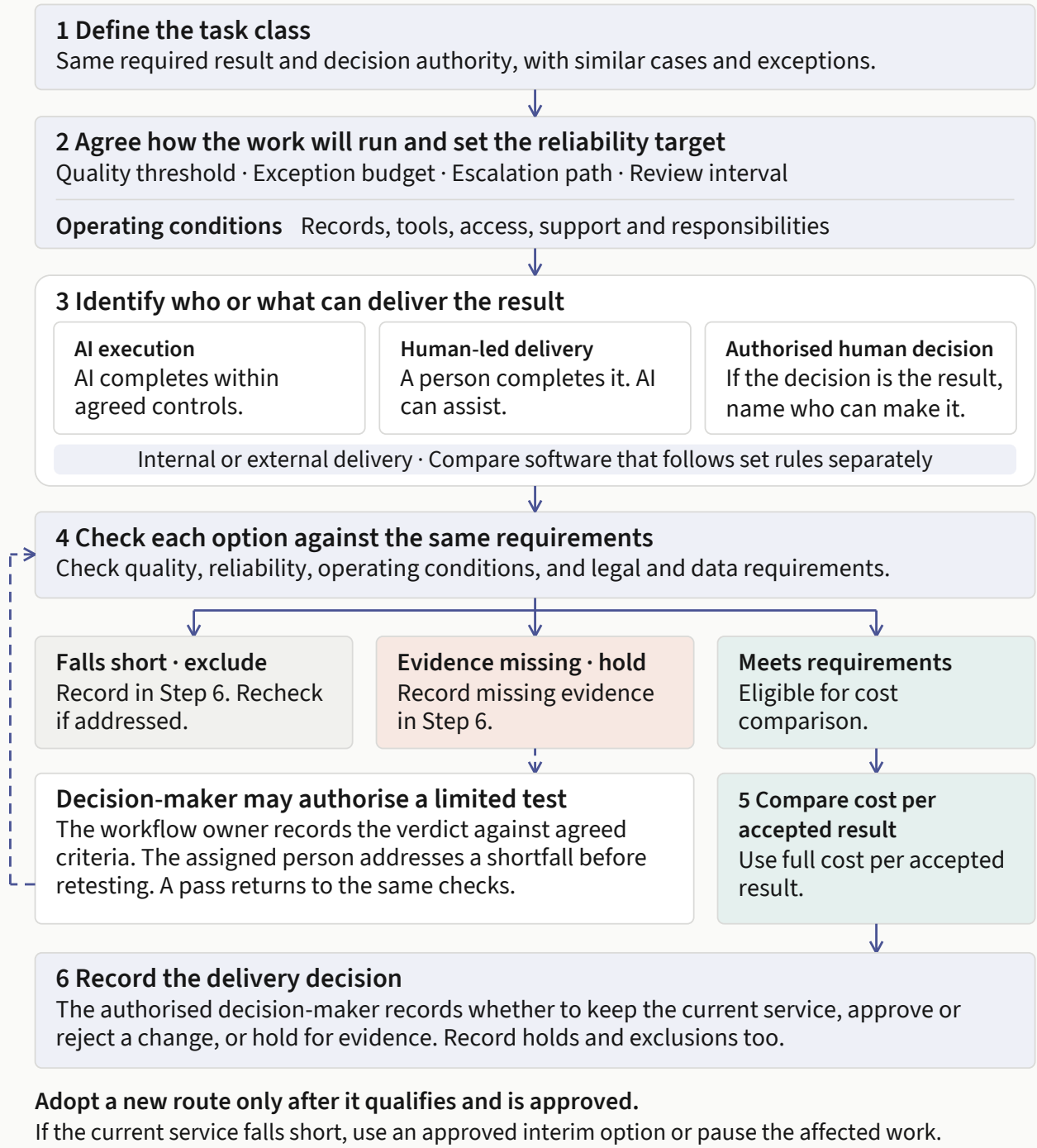


Figure 1. The flow follows one task class from the required result to the decision. The Northstar assessment example shows how a hold is documented.

© 2026 Radion Barakov · LIA

5. From assessment to decision

A cost comparison can make a different way of working look promising. Before deciding whether to try it, it helps to see how the service runs today, including what happens when a case takes more work than usual. The assessment gives the decision-maker three things.

1 · Cost baseline

What it costs to deliver the service today, including checks and corrections.

2 · Routing map

Different ways to deliver the same result, showing what people and AI would each do and who could provide the service.

3 · Workable option or reason to hold

Record which option the evidence supports and what it needs to run. If the evidence is incomplete, record what is missing and what would let the decision-maker reconsider.

An AI-assisted route might look promising on cost while the person verifying the results still needs access to some records. The decision-maker can hold off and reconsider a limited trial once that access is confirmed.

What the decision rests on

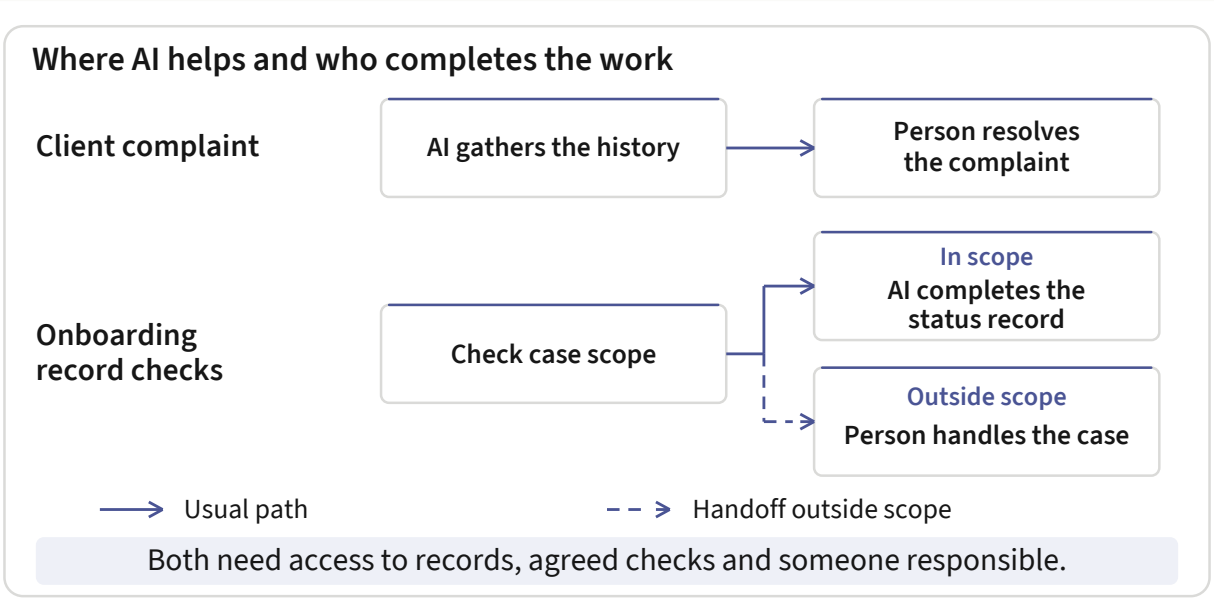
In that example, the **recommendation record** shows what access is missing and why the decision-maker is holding off on authorising a limited test. It separates what the assessor observed from what was estimated, so someone coming back to the decision can see what still needs checking.

The **LIA operator** applies the method and discloses any interest in supplying the service. The **assessor** prepares the recommendation and can also be the operator. If an independent review or procurement step is required, the record makes that clear.

The **authorised decision-maker** decides whether to proceed. The record shows what was decided, who decided it and when to review it. For a trial, it also says who can stop it and under what conditions. The **delivery provider** runs the agreed service, and contracting stays with the people authorised to do it.

The Northstar assessment example in Appendix C shows how the assessor documents a hold and who owns the next decision.

© 2026 Radion Barakov · LIA



The top row shows a planned handoff. AI gathers the history and a person resolves the complaint. In the bottom row, the workflow checks scope first. AI completes the status record within agreed controls for in-scope cases. The dashed arrow sends cases outside scope to a person. If a coordinator verifies every draft before completing the record, the coordinator still finishes the check.

How far has the work changed in practice?

The illustration shows how AI can help prepare a result or complete it within agreed controls. The **L0-L5 diagnostic** describes how each workflow operates. A function such as finance or client services can contain workflows at several levels. The diagnostic looks at how regularly AI is used, who completes the work, who can maintain or change the arrangement, and how much routine maintenance runs on its own.

Take the onboarding record check above. If the coordinator routinely uses AI to prepare a draft and still completes every check, that is repeatable use at **L2**. If the AI workflow begins to complete a defined class of status records in normal service, with a named owner and a way to refer cases outside scope, responsibility for completion has changed. That is the shift to **L3**.

The level gives the assessment a starting point. The decision about going further still depends on what the service needs, whether the proposed route can meet those requirements, and what it would take to run it well. Appendix B sets out all six levels and the evidence to look for. Northstar asks what another way of handling supplier cases would have to show before it could take them through to closure.

© 2026 Radion Barakov · LIA

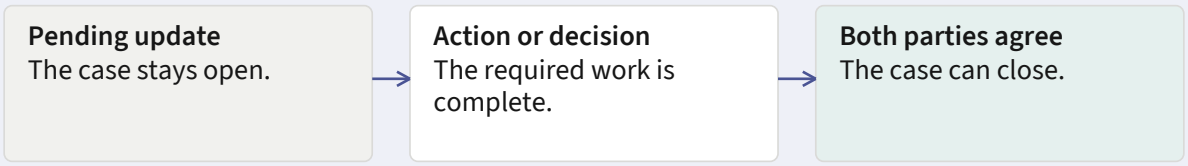
6. Northstar: who should close supplier cases?

When a supplier asks why an invoice is unpaid, someone in finance usually needs to look at the invoice, the purchase order and the payment record. Northstar's finance lead wants to know whether a qualified outside provider using AI could handle those cases through closure under the same agreed requirements. Northstar and the figures in this example are fictional.

What does the supplier need?

Several jobs can sit behind one unpaid invoice. Checking it against the order gives finance a record to work from. An update can tell the supplier what is pending and what happens next. The case stays open while the work is pending.

What counts as a closed case?



The Northstar assessment example shows what the owner captures.

Before closing a case, the handler obtains any required payment ruling from the finance lead and any required compliance sign-off from the authorised person. The case stays open until both parties agree to close and the handler records that agreement. A supplier reply confirming resolution can show the supplier's agreement.

Any agreed deadline is counted in business days and can vary by case. When both sides agree to an extension under the exception rules, the record keeps both deadlines, whether each was met, who approved the extension and what it cost.

Why is Northstar holding the transfer?

Suppose the invoice is in the system but the purchase order shows a different price. AI could pull up both records so the case handler can explain the discrepancy. If the amount is disputed, the finance lead decides whether to pay it. The case handler then follows through with the supplier.

That gives Northstar a picture of how the work could be shared. Before moving it, the finance lead needs to know how many cases come in, how often they close as agreed, what the current service costs and whether an alternative could meet the same standard. For now, the work stays where it is.

© 2026 Radion Barakov · LIA

What does Northstar need to check?

Record item	Northstar supplier cases
Quality threshold	Complete the required work and record both parties' agreement to close. A closure under an approved extension can count if it meets the revised deadline and the other acceptance criteria. Track the original deadline separately.
Options	Keep the current service, add AI support locally, or consider a qualified nearshore or offshore provider using the shared platform.
Full cost per accepted result	Establish incoming volume, accepted closures, delivery time, retained decisions, checking and rework.
Exceptions and authority	Finance lead sets the budget and review interval from the evidence. Record departures and incremental cost. Retain disputed-payment authority.
Interest in supplying the service	Record who could benefit from supplying or coordinating it.
Decision	**Hold the transfer.** Continue internal delivery where it meets the agreed requirements while checking closure, cost, access and responsibilities.

What would a limited test check?

In either AI-supported option, AI gathers the records and a case handler follows up with the supplier. The payment and compliance decisions stay with Northstar. What changes is who provides the service.

An estimate can show whether a limited test is worth exploring. Northstar would use evidence from the test to inform a separate decision on whether to transfer the work. Before a test can start, the finance lead needs to agree on the cases, acceptance criteria, stop conditions and review date, and confirm access, security, contracting and decision authority. The trial should include harder cases such as disputed payments and agreed deadline exceptions.

The workflow owner records whether the test met the agreed criteria. If it falls short, the assigned person addresses the shortfall before the option is tested again. A change to the service requires approval.

© 2026 Radion Barakov · LIA

Comparing the costs of three options

Suppose each option handles the same mix of 1,000 cases over 12 months. A case counts when it closes under the agreed criteria, including any agreed deadline.

To qualify, each option must reach at least 95% accepted closures and meet the operating conditions. The estimates are fictional. Northstar would also track how often it met the original deadline when one was agreed.

First-year cost per accepted closure

Delivery option	Base estimate	Saving vs today
Current internal service	€114.58	-
Northstar's team with AI	€102.04	€12.54 (10.9%)
Outside provider with AI	€87.63	€26.95 (23.5%)

Both AI-supported teams use Northstar's shared platform. Against Northstar's own team using AI, the provider costs an estimated 14.1% less per accepted closure. Under the cautious assumptions, the difference falls to 5.8%.

What goes into the first-year cost?

First-year cost component	Current internal	Local + AI	External + AI
Direct delivery	€72,000	€60,000	€30,000
Platform and tools	€3,000	€4,000	€4,000
Routine quality checks	€1,000	€2,000	€8,000
Oversight, access and retained authority	€22,000	€22,000	€24,000
Ongoing service coordination	€8,000	€4,000	€4,000
One-off implementation and transition	€0	€4,000	€6,000
Expected rework	€4,000	€4,000	€6,000
Proposed external test	€0	€0	€3,000
Full first-year cost	**€110,000**	**€100,000**	**€85,000**
Accepted closures out of 1,000	**960**	**980**	**970**

The two AI-supported options assume similarly skilled handlers and 1,000 paid hours each. Northstar uses €60 per in-house hour and €30 per provider hour, including the provider's margin. Today's service is modelled on 1,200 hours at the same in-house rate.

The shared workflow is assumed to reduce ongoing coordination. The external handover adds €2,000 to implementation and transition.

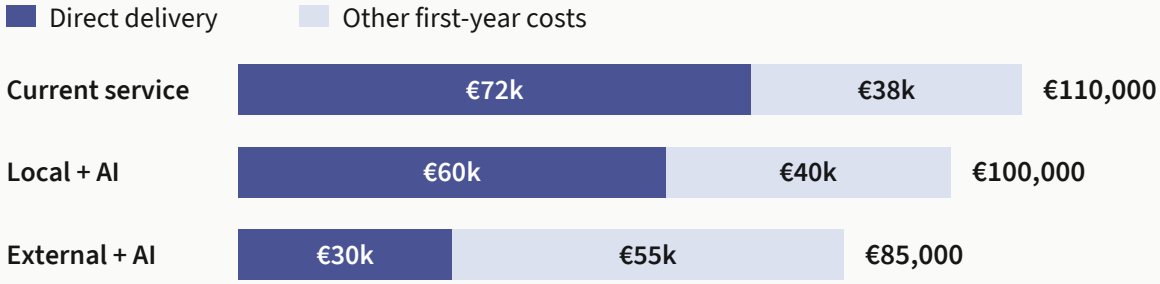
© 2026 Radion Barakov · LIA

Figure 2. Northstar decision and comparison

Hold the transfer **Internal delivery continues** where it meets the requirements — Decision-maker: finance lead

Evidence needed	Incoming cases · Accepted closures · Full cost
Conditions to check	Access and security · Contracting · Decision authority

What makes up Northstar's first-year cost?



Fictional base estimates · Same 1,000 cases · External option includes its proposed test

The pale part of the provider's bar includes Northstar's time. If two prices disagree, AI can find the invoice and the order. Finance decides what to pay, and the handler checks back with the supplier so they can agree the case is settled.

What changes in the cautious case?

Local handling time falls 3% and external handling time rises 6%. That puts direct delivery at €58,200 and €31,800 respectively. Oversight and rework each fall by €1,000 locally and rise by €1,000 externally. The provider closes ten fewer cases to the agreed standard.

Cautious first year	Local + AI	External + AI
Full cost, including the external test where applicable	€96,200	€88,800
Accepted closures	980	960
Cost per accepted closure	**€98.16**	**€92.50**

Both estimates clear the 95% threshold. Northstar would still hold the transfer if too few disputed cases had been tested.

Northstar would compare ongoing delivery, support and maintenance separately from implementation and the one-off test, including work its own team still has to do.

© 2026 Radion Barakov · LIA

Five years of supplier cases at Northstar

Let’s see what more cases cost at year-one prices, with the same mix of queries and disputes.

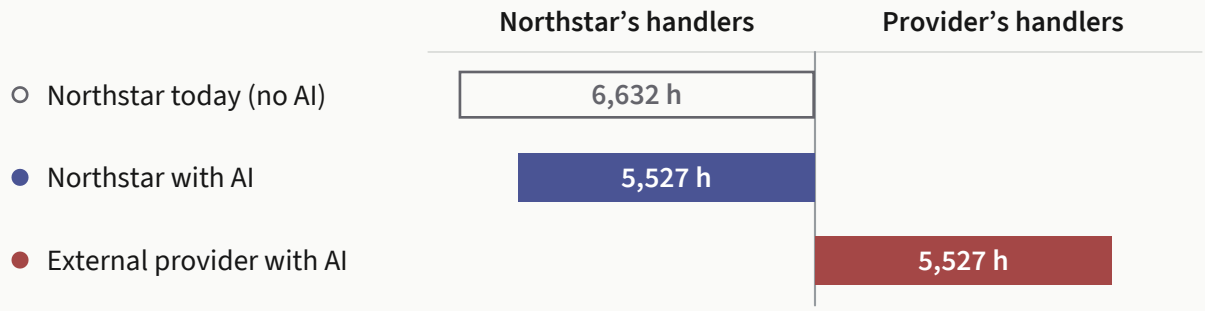
1,000 cases in year 1
Supplier cases received

5% more cases a year
1,216 cases in year 5

12 min less per case
72 min today, 60 with AI

Who would handle the cases?

Hours of direct case handling over five years.



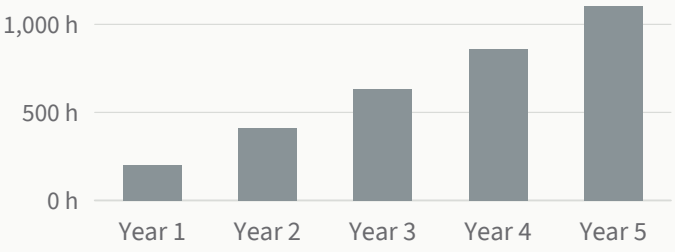
The full cost covers Northstar’s time to check the provider’s work and decide what to pay in a dispute.

Service	Cases resolved	€ per resolved case	Full cost	Saved vs today
Northstar today (no AI)	5,306	€114.58	€607,970	0%
Northstar with AI	5,417	€98.69	€534,592	12.1%
External provider with AI	5,362	€80.02	€429,052	29.4%

Compared with Northstar’s team using AI, the provider could close nearly as many cases for less. Assume ongoing costs rise 5% a year. Setup and testing are counted once.

Direct handling hours saved so far

Either AI team saves 200 hours in year one on supplier cases alone.



A second service in year two?

Try sharing the work on monthly reports. If that frees time too, test a third service.

If you ran Northstar, where would you put those hours to work?

© 2026 Radion Barakov · LIA

7. Fableoak: sharing the work behind monthly reports

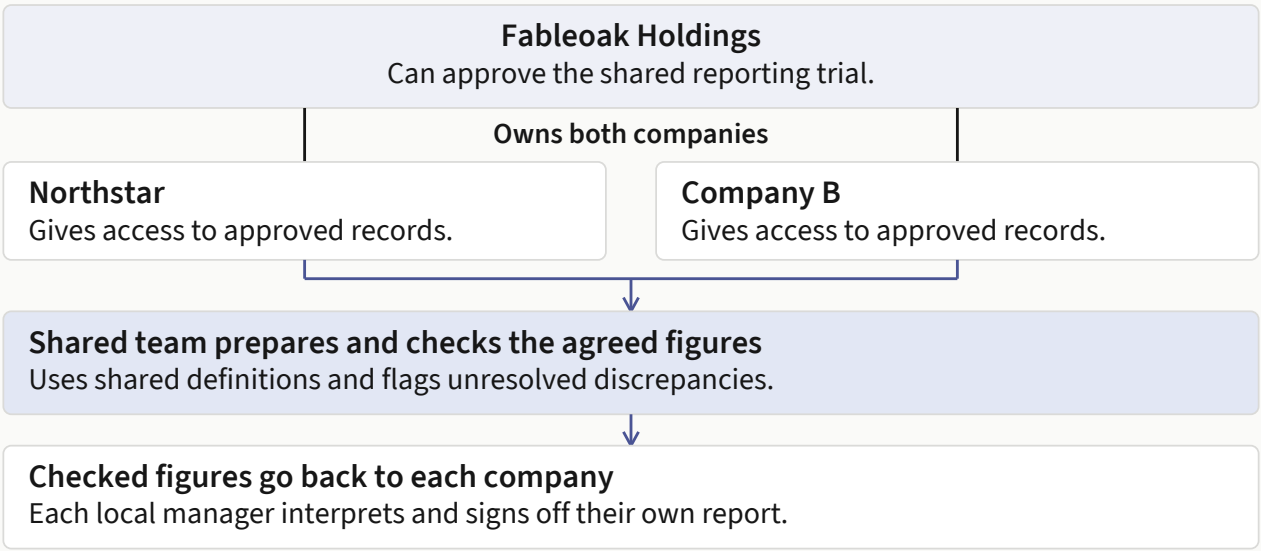
Fableoak Holdings, the fictional owner of Northstar and several related service businesses, wants to share some of the work behind their monthly reports.

At month-end, Northstar counts a reopened client case as open as soon as it comes back. Company B counts it only when someone starts working on it again. Both reports can say “open cases” and give a different picture of how much work is waiting.

A shared service could start with figures the companies already define the same way. It could prepare a set of checked figures by the deadline, with their sources and any unresolved discrepancies. The trial would show whether the shared service can deliver those figures to the agreed standard. Fableoak can then decide where AI would help.

- **Which figures could be shared?** Recent reports show where definitions match and whether the records can be checked.
- **What would stay with each company?** Local managers would still interpret the figures and sign off their reports.
- **What would it cost each company?** The comparison includes each company’s share of setup, checks and corrections, and counts a set of figures only if it meets the agreed standard by the deadline.

Fableoak’s shared reporting trial



© 2026 Radion Barakov · LIA

When would Fableoak extend the service?

The first trial could cover a few figures that two companies already define the same way. Fableoak can approve the shared trial, while each company decides which records the shared service can see and names who can stop it. Anyone recommending the trial who could earn money from running the service discloses that interest.

Suppose both companies agree how to count reopened cases before the trial. Northstar gets usable figures on time, but Company B spends hours reconciling case records that still reflect its old rule. That time belongs in Company B's cost. The set counts only if the corrected figures meet the agreed standard by the deadline. Before extending the service, Fableoak looks at the results, costs and time freed up for each company. A group average could hide where the work has shifted.

Suppose both sets of figures now meet their company's standard. One cycle is a month's reporting run. The fictional costs compare separate and shared preparation, including checks and corrections. The monthly figures include each company's allocated share of setup. In practice, Fableoak would record how many months it spreads setup across and how it assigns each share. Before any further transition work, the group estimate is 185, a 7.5% saving.

Company	Meets standard?	Cost units per cycle	Decision and support
Northstar	Yes	100 to 65	Retain local sign-off under shared preparation.
Company B	Yes	100 to 120	Agree compensation and assign an owner to reduce the added coordination.
Group	**Yes, both**	**200 to 185**	**Agree and cost the group-funded transition support for Company B first. Extend if both companies meet their agreed criteria and the revised group saving meets the agreed minimum.**

Compensation shifts who bears the cost. Extra staff time or tools change the group's full cost. Hold a rollout that fails a company's standard and recheck the economics if only one company proceeds.

The diagram on the preceding page shows where the shared work ends. The shared service prepares and checks figures the companies count the same way. Each company still interprets its own report and decides what to do. When another company joins, a recent report gives Fableoak a starting point for seeing which figures could be shared, what it would take to connect the records and who would resolve discrepancies.

© 2026 Radion Barakov · LIA

8. Testing LIA: evidence and limits

What can a trial tell us?

Fableoak's trial can show whether each company gets a usable set of checked figures on time and at the expected cost. To see whether LIA helps people make better decisions, we'd need to compare similar decisions made with and without it.

Dillon and colleagues (2025) studied access to an AI tool across 66 firms over six months. Workers assigned access spent less time on email. The researchers detected no change in the number or mix of tasks recorded in the applications they measured. The study shows why it's worth looking beyond the time saved on one activity. To judge a service, a business also needs to see what people received, what it cost to deliver and how any time freed up was used.

At Fableoak, the checked figures might arrive sooner while local staff still have to check figures the shared service was meant to verify. That extra work belongs in the cost comparison. If the figures fall short of the agreed standard, extra costs exceed the exception budget or the advantage drops below the agreed minimum, Fableoak has a reason to revisit the recommendation. The record also needs to show how the original choice followed from the evidence and assumptions.

Table 3. What the evidence can tell you.

When	What you can learn	What's still open
Before a change	What each option would need and whether there's enough evidence to proceed.	How the proposed service will perform.
After a change	Whether the service met the standard and the expected cost advantage held.	What explains the outcome.
Across comparable decisions	Whether LIA helps assessors make better-supported recommendations and more accurate forecasts.	How well the finding carries over to other work.

A test of LIA would compare decisions about similar work under similar conditions. It would record what each business expected before it acted, then assess what happened using the same measures of quality and cost. That comparison has yet to be carried out.

© 2026 Radion Barakov · LIA

What published studies measured

Study	Finding	Evidence limit
Brynjolfsson, Li and Raymond (2025)	Across 5,172 support agents, AI assistance increased issues resolved per hour by 15% on average. Effects varied by experience and skill.	One company, service setting and tool.
Dell'Acqua and colleagues (2026)	GPT-4 helped on several consulting tasks. On the brand choice, the AI groups were about 19 percentage points less likely to answer correctly than the control group.	One brand task in an experiment involving 758 consultants overall.
Dillon and colleagues (2025), preprint	Across 66 firms and 7,137 workers, assigned access reduced email time by 1.4 hours per week in months four to six. The study detected no change in measured task mix.	Application activity was measured. Business performance was outside the study.

The ILO review by Merola and colleagues (2026) discusses how productivity and work organisation vary across settings. These studies inform what to measure when using LIA. Appendix D explains what each source can support.

How would you test the method?

We could compare similar delivery decisions made with LIA and with the organisation's usual way of deciding how work gets done. Both approaches would start with comparable evidence, including routine and difficult cases. Reviewers could then judge the recommendations, the quality and cost forecasts and the time each assessment took.

Where a proposed route goes ahead, a comparison group or matched cases could help show what changed in accepted results, full cost and unplanned corrections. Track how long decisions stay on hold and what waiting costs where it can be measured. Judge the hold on the evidence available at the time. An approved trial can help test the reason for holding. These evaluations have yet to be carried out.

Testing whether LIA improves delivery decisions

For example, should a business change who handles its supplier cases?

Start with comparable decisions
Use LIA and the usual process with comparable evidence.

→

Review the recommendations
Check reasoning, quality and cost forecasts, and assessment time.

→

Follow up on decisions
Check results and costs after changes. Revisit holds as evidence comes in.

© 2026 Radion Barakov · LIA

9. How LIA fits with existing approaches

Before Fableoak can share the work behind its monthly reports, it needs to know how each company prepares them, what that work costs and who can access the records. Process mapping, costing, governance and procurement each help answer part of that decision.

There's also a question about what changes when AI helps with one part of the work. Wu and colleagues (2026), writing as the CIVIC-AI Collaboration, propose looking at the whole workflow to see where value lasts, where people retain control and what they have room to learn. For Fableoak, the point is practical. Quicker preparation matters only in the context of the whole reporting process, including what local managers still have to check, interpret and act on.

Table 4. LIA alongside existing approaches.

Approach	Main question	Fit and limit
APQC · Process classification	Where does the work sit?	**Fit.** Both distinguish kinds of work. **Limit.** APQC maps processes. LIA compares ways to deliver a specific result.
Fügener et al. · Human-AI task allocation	Who should do which tasks?	**Fit.** Both consider what people and AI do well. **Limit.** LIA also considers service cost and decision authority.
CIVIC-AI · AI across a workflow	What changes across the workflow?	**Fit.** Both look at how work shifts and where people retain control. **Limit.** CIVIC-AI examines how AI shapes a workflow over time. LIA compares delivery options for an agreed result.
Chen et al. · Model routing	Which model should handle the request?	**Fit.** Both consider capability and cost. **Limit.** Model routing selects a model. LIA compares ways to deliver the service.

© 2026 Radion Barakov · LIA

Table 4. LIA alongside existing approaches (continued).

Approach	Main question	Fit and limit
Kaplan and Anderson · Activity-based costing	What does the work cost?	**Fit.** Both account for the time and resources work takes. **Limit.** Costing estimates what delivery costs. LIA also asks whether an option meets the service requirements.
NIST · AI Risk Management Framework	How should AI risk be managed?	**Fit.** Both consider risk in deciding whether a change can proceed. **Limit.** NIST covers AI risk more broadly. LIA focuses on one delivery decision.
ISO/IEC 42001 · AI management systems	How should an organisation manage AI?	**Fit.** Both give attention to oversight and responsibilities. **Limit.** ISO/IEC 42001 addresses the organisation's management system. LIA addresses a specific delivery choice.
CIPS · Sourcing and procurement	Should we buy the service, and on what terms?	**Fit.** Both can inform a choice of provider. **Limit.** Procurement covers supplier selection and contracting. LIA also compares how the work could be done.

For Fableoak, the process map helps locate the reporting work, costing shows what it takes today, and governance sets the conditions for sharing records. A make-or-buy assessment can already compare costs and service standards. LIA spells out who completes each kind of work, where AI helps and who can make decisions. Would sharing the preparation give each company checked figures it can use in its own report, at a cost that makes the change worthwhile?



Ask two people involved in a recurring job what counts as finished for the person receiving the result. Agree on that before comparing ways to do it.

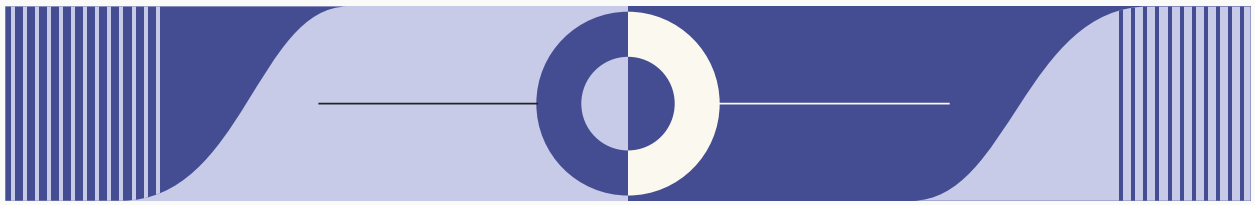
© 2026 Radion Barakov · LIA

Closing note

While I was writing this paper, I watched AI go from something people tried on their own to something companies had to write policies for. Some businesses reorganised while people were still figuring out how to use it in their jobs. I saw people get real help from AI and still have to work out for themselves how it fitted into the way the service was delivered. That helped me understand why making one part of the job easier doesn't always change what the business delivers. The longer-term difference depends on how the business supports people through the change and makes clear who is responsible for the work.

If I were looking at a service with a business owner, I'd start by getting the whole picture. I'd want to hear why the business provides the service, what it needs to deliver, what other ways of doing it are open to it and who is responsible at each step. Then I'd follow a few real cases from request to result and check that picture against what actually happens. That would give us a better basis for deciding what's worth testing. Appendix C sets out what I'd record along the way.

I began by asking whether a different way of working could still deliver a useful result. I'd look at what people can do with the result they receive. If the service holds up, where does any saving come from, and does it last once the work becomes routine? I'd like to find out by following staff time, provider fees and the work the business still has to do itself.



© 2026 Radion Barakov · LIA

Appendix A. Glossary

Term	Meaning in this paper
Labour-Intelligence Arbitrage (LIA)	A method for comparing the full costs of ways to deliver the same recurring result under agreed standards and operating conditions.
Arbitrage	The possible cost advantage between arrangements that deliver the same result to the same standard.
Routing and route	Routing is choosing how the work gets done. A route includes who completes it, the steps, support and checks.
Sourcing	Who provides the work. Employees, independent specialists or people working for a provider can do all or part of it. Any of these arrangements can use AI.
Operating conditions	The records, tools, access, support and responsibilities people and systems need to do the work.
Function	An area of business work that can span teams, such as client services.
Task class	Recurring work with the same required result, similar cases and exceptions, and the same rules about who can decide.
Completion destinations	For routes involving people or AI, who or what delivers the result: AI execution, human-led delivery or an authorised human decision.
AI execution	An AI workflow delivers the required result within agreed controls, with a named person accountable for its use.

© 2026 Radion Barakov · LIA

Term	Meaning in this paper
Human-led delivery	A person delivers the required result, with or without AI support.
Authorised human decision	A named person makes and owns the decision when the decision itself is the required result.
Lowest full cost per accepted result	The starting point for comparing options that meet the agreed standard and operating conditions. A decision-maker can choose a higher-cost option for a specific benefit and record why.
Full cost per accepted result	Divide full cost, including failed attempts and rework, by the number of results meeting the agreed standard.
Required result	What the completed work must deliver.
Quality threshold	What a result must achieve to count as accepted, including any deadline.
Exception budget	How much extra cost the business can absorb from unplanned departures over a stated period. Planned handoffs count as normal cost.
Escalation path	Who takes a case that needs more expertise or authority, and why it goes to them.
Review interval	How often the arrangement is reviewed. A material change or breach may call for earlier review.

© 2026 Radion Barakov · LIA

Term	Meaning in this paper
Hold a proposed change	Pause the change until evidence or conditions are in place. If the current service falls short while a replacement still needs testing, use an approved interim arrangement that meets the requirements or pause the affected work. Any required stop rule still applies.
Limited test	A trial with safeguards and a defined scope to gather evidence for a later decision.
Assessor	The person or team that examines the work, compares options and prepares the recommendation.
LIA operator	The person or team applying LIA. They disclose any interest in supplying the service.
Delivery provider	The person, team or organisation responsible for the service it agrees to deliver.
Authorised decision-maker	The named person or group that can retain the current arrangement, approve or reject a change, hold for evidence or authorise a limited test.
Recommendation record	The evidence, options, costs, assumptions, interests and decision rights behind the recommendation.
Three assessment outputs	The **cost baseline** records current cost. The **routing map** compares ways to deliver the result. The third output is a **workable option or reason to hold**. Record what the supported option needs to run, or what's missing and what would let the decision-maker reconsider.
L0-L5 diagnostic	The L0-L5 diagnostic describes how AI has changed an individual workflow. A function can contain workflows at different levels.
Accepted result	A result that meets the agreed criteria, including a deadline or sequence where applicable.
Workflow owner	The named person responsible for the workflow and its test verdict.
Maintenance loop	The recurring process that runs the approved work, reviews triggers and refers proposed changes for approval.

© 2026 Radion Barakov · LIA

Appendix B. L0–L5 diagnostic

Reading the levels

Use the levels to see how AI has changed the work. Then compare each workflow's current and required level before deciding where to go further.

Finding the level in practice

The levels build from L0. For example, L4 adds organisational control and handoff readiness to the change in completion responsibility at L3. Look for evidence in normal service of how often AI is used, who completes the work, who controls it and what runs on its own.

Level	How the work runs	What you'd see
L0 · No AI	AI does not affect how the workflow runs.	The work runs without AI involvement that changes how it gets done.
L1 · Individual use	People use AI on their own, as needed.	Use is ad hoc, with no shared routine or change in responsibilities.
L2 · Repeatable use	People use AI routinely.	The same role remains responsible for completing the result.
L3 · Changed responsibilities	AI changes who or what completes the result in normal service.	There is a named owner and a way to handle exceptions.
L4 · Controlled delivery	A tailored, organisation-controlled route is the governed default.	The organisation demonstrates that it can maintain it, hand it to a replacement or stop it.
L5 · Self-running maintenance	Routine routing, exception handling and review run within approved rules without human intervention.	The loop responds to a trigger. The owner approves changes and can stop it.

At L4, a tailored route becomes part of running the function. The organisation controls its instructions, connections, settings and data. It can maintain the arrangement or change providers, and has named oversight and stop authority.

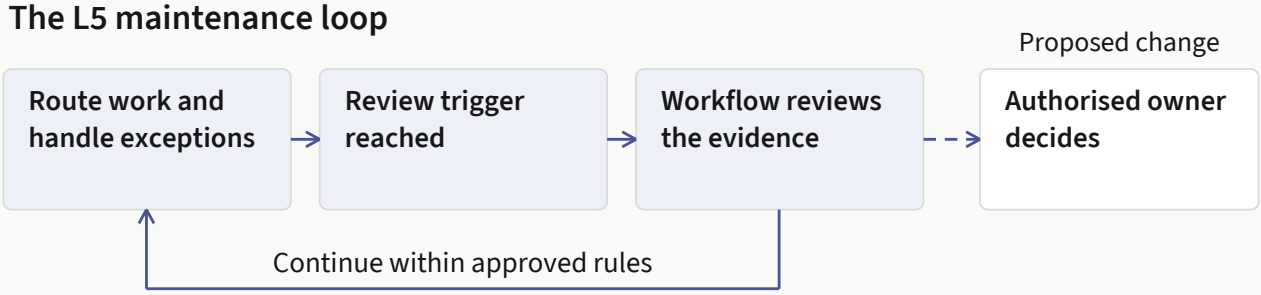
© 2026 Radion Barakov · LIA

What changes between levels

A gap exists when the work needs a higher capability and the workflow has yet to show it.

Transition	What changes	What you'd see
AI usage gap · L2 to L3	AI changes who or what completes the result.	A deliberately chosen route delivers results during normal service.
AI governance gap · L3 to L4	The organisation controls the default way the work runs.	The organisation shows it can maintain the workflow and hand it to a replacement.
Maintenance Autonomy · L4 to L5	Routine routing, exception handling and review run without human intervention.	The loop runs within approved rules. A proposed change goes to the authorised owner. A named person can stop it.

The L5 maintenance loop



A change in workload or correction costs can bring a review forward.

One function, several levels

For an AI route, set the level the result, controls, oversight and handoff require.

Finance workflow	Current	Required	What to do
Routine invoice matching	L2	L3	Routine matches must clear without manual sign-off. Refer mismatches to Finance.
Supplier-case preparation	L3	L3	Retain while it meets the service requirements.
Month-end record checks	L4	L3	Keep the instructions and access needed for a handoff up to date.

A workflow can stay at its current level if it meets the service requirements. When maintenance runs on its own, people retain the decisions assigned to them.

© 2026 Radion Barakov · LIA

Appendix C. Assessment outline

Start with one service the business already provides. Set out what it needs to deliver, how it runs today and who can decide whether to change it. The Northstar assessment example applies the steps to a proposed transfer.

Step	Question to answer	What to record
1. Frame	Is this work worth assessing?	A clear question and named decision-maker, or a reason to stop.
2. Describe	How does the work run today?	Task classes, cost baseline and the four controls.
3. Compare	Which routes have shown they meet the standard and operating conditions, and how do their full costs compare?	A routing map of the current service and alternatives, with full costs for options that meet the requirements and evidence gaps marked separately.
4. Check feasibility	What needs to be in place before a test or change?	A workable option or reason to hold, recorded in the recommendation.
5. Decide	What should happen next?	A recorded decision, a review interval and any conditions for proceeding.

Estimates can help decide whether to run a limited test. Put safeguards such as approved data access in place before testing, then use the evidence to decide whether to adopt the arrangement. Record any evidence gaps and what would resolve them. If the current service falls short and the replacement still needs testing, the decision-maker can use an approved interim arrangement that meets the requirements or pause the affected work.

What to check after changing a service

Subject	Check in practice
Quality threshold	Did the result meet the agreed standard?
Full cost per accepted result	Divide full cost, including failed attempts and rework, by the number of accepted results.
People's time	How much time was released? What capacity became usable? Did spending fall?
Exceptions and extra cost	What unplanned work came up, and did its cost stay within the exception budget?
Named responsibility	Who runs the service, who can stop it and when will they review it?

© 2026 Radion Barakov · LIA

Fictional assessment

Northstar assessment example

The estimate gives Northstar a reason to look more closely. For now, the finance lead would hold the transfer and check how the current service performs and what access an outside provider would need. If a limited test is approved, its results go back through the requirements check. A route that qualifies can then be compared on full cost per accepted closure before the finance lead makes a separate transfer decision.

Record item	Northstar supplier cases
Question and owner	Should Northstar transfer supplier case handling? The finance lead owns the workflow and the decision and retains disputed-payment authority.
Required result	Complete the work, obtain any required payment ruling and compliance sign-off, and record both parties' agreement to close. Record approved exceptions and track the original agreed deadline separately.
What to measure now	Measure incoming cases, accepted closures, waiting, active work, retained Finance time and actual rework. The case figures remain assumptions pending that evidence.
Routing map	Compare the current service, local + AI and qualified external + AI. The AI-supported routes use the same platform. Northstar owns the business context and data and controls access.
Proposed test and criteria	Before approval, agree which cases count, the case mix, the common follow-up window, acceptance criteria, review date, exception budget and stop rules. Classify cases and set deadlines by the same rules across routes. Finance checks a sample against the records, including cases marked routine. Divide accepted closures by all included cases, even those open at the cutoff. Test disputes, deadline exceptions and busy periods. Hold the transfer even at 95% overall if too few disputes were tested. The finance lead records pass, shortfall or need for more evidence. Fix shortfalls before retesting.
Cost and interests	Verify platform allocation, checking, retained authority and rework. Set the minimum worthwhile advantage and record any interest in supplying the service.
Decision and next review	**Hold the transfer.** Continue internal delivery where it meets the requirements. Review after baseline and access checks, or earlier if a service or authority breach requires action. Any transfer requires approval.

© 2026 Radion Barakov · LIA

Appendix D. Evidence and source notes

The table shows where the paper draws on research, frameworks and guidance, and what each source can support.

Source	How the paper uses it	What it can show
Statistics Netherlands (2026), *Staff shortages mean business is turning to automation*	Shows why Dutch business-services firms are looking at automation.	Firms' reported responses to staffing shortages. The automation category includes robots and AI support.
U.S. Census Bureau (2026), Stepler, *AI use at work*, March survey	Puts reported time savings in everyday work into context.	Workers' own estimates of time saved, among those who used AI at work in the previous week.
Brynjolfsson, Li and Raymond (2025), *Generative AI at Work*, QJE 140(2), 889-942	Shows how effects vary across support workers.	One company, occupation and tool.
Dell'Acqua et al. (2026), *Navigating the Jagged Technological Frontier*, Organization Science 37(2), 403-423	Shows how effects vary across consulting tasks.	Specific tasks, participants and model.
Dillon, Jaffe, Immorlica and Stanton (2025), *Shifting Work Patterns with Generative AI*, arXiv:2504.11436v4	Separates email time from measured work.	Application activity, with business performance unmeasured.
Wu and colleagues (2026), writing as the CIVIC-AI Collaboration, *When Does AI Augment Work? A Workflow-Level Framework for Human-Agent Collaboration*, arXiv:2609.12482v1	Frames what changes across a workflow.	A proposed framework with a worked example.
Merola et al. (2026), ILO review of GenAI, jobs, productivity and work organisation	Frames questions about released time and effects on work.	A review across varied settings.

© 2026 Radion Barakov · LIA

Evidence and source notes (continued)

Source	How the paper uses it	What it can show
NIST (2023), AI RMF 1.0 Core	Provides context on AI risk.	Guidance on managing AI risk.
ISO/IEC 42001:2023	Provides context on AI management systems.	A standard. Compliance needs separate evidence.
Fügener, Walzner and Gupta (2026), *Roles of Artificial Intelligence in Collaboration with Humans*, Management Science 72(1), 538-557	Informs the human and AI roles comparison.	Image-classification experiment and simulation.
Chen, Zaharia and Zou (2023), *FrugalGPT: How to Use Large Language Models While Reducing Cost and Improving Performance*, arXiv:2305.05176v1	Informs model choice within a route.	Model-level cost and quality.
APQC (n.d.), Process Classification Framework, online guidance	Shows how recurring work can be grouped.	A classification framework.
Kaplan and Anderson (2003), *Time-Driven Activity-Based Costing*, HBS Working Paper 04-045	Informs the cost baseline.	A method for estimating costs from time and capacity.
CIPS (n.d.), *The thirteen steps of the procurement process*	Informs make-or-buy and supplier decisions.	Procurement guidance.

The examples are fictional. LIA's effectiveness in practice remains to be tested.

Icons: Lucide contributors and Phosphor Icons.

© 2026 Radion Barakov · LIA