

TRUST ASSESSMENT / SUPPORTED CONCLUSIONS

# HOW TRUSTWORTHY IS OUR ANALYSIS

The question is how organizations should adopt artificial intelligence (AI) to improve productivity and profit while protecting information, accuracy and continuity. Bellhaven seeks \$250,000 in additional annual profit attributable to AI within 24 months, with a \$120,000 first-year cash ceiling and an initial decision within 45 days.

WHAT THE CURRENT FINDINGS JUSTIFY

The analysis provides a dependable method for selecting and evaluating AI uses, and its principal financial calculations reconcile. Bellhaven has credible opportunities in bookkeeping and tax processing that merit limited tests. Realized staffing savings, a fully costed annual return and readiness for wider use remain unestablished.

## The Trust Levels Describe Support for Each Conclusion

- Well supported.** Strong support within the stated scope; remaining uncertainty is unlikely to change the conclusion materially.
  - Partly supported.** Meaningful support exists; an unresolved issue could materially change the conclusion.
  - Not established.** Available information or completed examination is insufficient for reliance.
  - Contradicted.** Available information directly conflicts with the conclusion. Marked CONFLICT.
- Colors describe support, including for adverse findings. The levels are not probabilities or an overall score.

| CONCLUSION ASSESSED                                            | TRUST LEVEL      | REASON FOR THE RATING                                                                                                         |
|----------------------------------------------------------------|------------------|-------------------------------------------------------------------------------------------------------------------------------|
| The method covers the main AI adoption decisions.              | Well supported   | Selection, safeguards, staff adoption, cost, attribution and continuing review receive substantive treatment.                 |
| Bookkeeping is a justified first candidate for a limited test. | Partly supported | The workload gives a measurable route to avoided hiring. Client demand, usable time savings and test costs need confirmation. |
| The stated financial scenarios follow from their assumptions.  | Well supported   | Recomputation confirms the principal staffing, tax, recurring-profit and cash calculations.                                   |
| The proposed AI uses will deliver the annual profit target.    | Not established  | The favorable modeled return is below the target, and some continuing costs and operating conditions remain unpriced.         |
| Bellhaven is ready for wider deployment.                       | Not established  | No live performance, implemented safeguards or confirmed resource assignments were available for examination.                 |

TRUST ASSESSMENT / QUESTION COVERAGE

# The Analysis Addresses the Full AI Adoption Question

The guidance considers business owners, managers and operating teams across industries. Bellhaven supplies a concrete accounting application; the wider method also addresses retail, manufacturing, software, sales, internal functions and public services.

| WHAT THE QUESTION REQUIRES             | WHAT THE ANALYSIS ESTABLISHES                                                                                            | WHAT REMAINS UNRESOLVED                                                                                                     |
|----------------------------------------|--------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------|
| Choose where to begin.                 | Compares targeted AI uses, existing software, ordinary automation, staffing, process improvements and waiting.           | Bookkeeping has a plausible financial route, but no completed comparison proves it is Bellhaven's highest-value first test. |
| Protect information and accurate work. | Specifies access restrictions, qualified review, controlled actions, error tests and recovery procedures.                | Vendor terms and the actual configuration have not been tested or approved in this examination.                             |
| Encourage effective staff adoption.    | Distinguishes unclear permission, lack of skill and unsuitable tasks; measures accepted work and total effort.           | The survey suggests barriers. Training effectiveness and managers' ability to reassign work remain unmeasured.              |
| Measure profit attributable to AI.     | Separates hours, avoided spending, additional contribution and program costs; explains comparisons with work without AI. | Bellhaven still needs an executable comparison and financial records linking saved hours to changed spending.               |
| Control costs and stay current.        | Includes staged spending, supplier alternatives, monthly monitoring, quarterly review and retesting after changes.       | Actual maintenance costs, available staff time and responsibility for ongoing work need confirmation.                       |

## Research Supports Local Testing Without Predicting Local Profit

Direct examination of accounting research confirms the distinction between an association with more clients supported and an experiment finding better transaction classification. Those results support testing accounting tasks; neither establishes Bellhaven's annual profit.

A workplace experiment distinguishes access to AI from induced use. A developer-study update reports selection problems that weaken its effect estimate. The analysis appropriately avoids converting either finding into a universal productivity percentage.

THE COVERAGE IS STRONGER THAN THE LOCAL FORECAST

Identifying useful opportunities and decision rules answers a substantial part of the question. Uncertain implementation outcomes do not invalidate that work. They determine how far Bellhaven can rely on the projected financial benefits.

# Bookkeeping Savings Must Remove the Need for Added Staff

The supplied workload is 11 hours per bookkeeping client each month, including 6.5 sorting hours. Twenty additional clients require 2,640 hours annually. At 200 clients, sorting accounts for 15,600 annual hours. Each planned position costs \$98,000 and supplies 1,400 client hours.

| RECOMPUTED FINDING                                                             | PRACTICAL MEANING                                                                          | CONDITION FOR RELIANCE                                                                                 |
|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| Avoiding one position requires 1,240 hours saved, about 7.95% of sorting time. | One position remains to cover up to 1,400 hours of additional work.                        | Both positions must otherwise be necessary; quality and other work must remain acceptable.             |
| Avoiding both positions requires 2,640 hours saved, about 16.92%.              | At 20%, savings are 3,120 hours, leaving 480 after the new workload.                       | Savings must occur in the right teams, roles and months. Annual totals alone cannot establish this.    |
| A 0.4-hour monthly saving per client provides 960 annual hours at 200 clients. | The remaining 1,680 hours still require two positions under the supplied staffing rule.    | The 0.4-hour result is a calculation from supplied test values; it is not verified live performance.   |
| At eight additional clients, workload rises by 1,056 hours.                    | At 188 clients, about 7.20% sorting improvement could avoid one genuinely needed position. | Only eight clients are signed. The hiring plan without AI must be confirmed before recording a saving. |

## The Comparison Correctly Deducts Improvement Without AI

The supplied pilot-team reductions average 0.9 hours; comparison-team reductions average 0.3. Deducting the 0.3-hour comparison improvement and 0.2 hours of added review leaves 0.4 hours, about 6.15% of initial sorting time. The arithmetic supports the conclusion that these values do not justify expansion on avoided-hiring grounds.

## The Quality Results Cannot Establish Safe Wider Use

The supplied error averages are 0.8 per file with AI and 0.6 without it. Sample sizes, severity and variation are insufficient to establish the size or cause of that difference. Catching four of four planted critical errors and 13 of 16 minor errors identifies test performance, but cannot establish reliable detection across future work.

A POSITIVE TIME SAVING CAN STILL FAIL THE BUSINESS TEST

Four teams provide few independent comparisons. The pairwise net range of 0.1 to 0.7 hours describes variation in the supplied values; it is not a statistical confidence interval. Expansion needs acceptable quality and usable savings large enough to change a real spending decision.

# Tax Savings Depend on Contractor Work and Senior Review

Contribution means revenue after additional delivery costs. The tax benefit depends on whose hours disappear. The supplied 2,600 returns require 2,860 entry hours, including 1,700 contractor hours. A 50% entry reduction is an unverified vendor performance assumption. Additional verification requires 650 senior-accountant hours.

| ASSUMPTION TESTED                              | RECOMPUTED FINANCIAL RESULT                                                                                                                  | AUDIT JUDGMENT                                                                                                                          |
|------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| Entry savings remove contractor hours first.   | Avoided spending is \$100,100. After \$19,600 first-year fees, the gain is \$80,500; after \$15,600 recurring fees, it is \$84,500 annually. | The arithmetic is well supported. Allocating all 1,430 saved hours to contractor work requires operational confirmation.                |
| Entry savings are proportional across staff.   | Avoided contractor spending is \$59,500. The first-year gain is \$39,900 and recurring gain is \$43,900.                                     | This is a separate allocation assumption. Salaried time saved does not automatically reduce payroll.                                    |
| Verification displaces other paid senior work. | At \$175 contribution per hour, 650 hours cost \$113,750 in forgone contribution. First-year results become losses of \$33,250 or \$73,850.  | The loss applies only when valuable work is actually displaced. First-year gains disappear at 460 or 228 displaced hours, respectively. |
| Entry improvement reaches 35%.                 | First-year gains before displaced senior work are \$50,470 with contractor-first savings or \$22,050 with proportional savings.              | The opportunity can remain positive below the vendor assumption, but review demands can still erase the benefit.                        |

## A Seasonal Reviewer Is a Credible Competing Use of Resources

The supplied declined work offers \$187,000 potential revenue. After 45% delivery costs and a reviewer costing \$49,400, the calculated contribution is \$53,450. About \$89,818 of additional revenue covers the reviewer at the assumed 55% contribution margin. Customer availability, preparation capacity and sufficient review hours remain unconfirmed.

Review accounted for 34 of 60 delayed files and missing client documents for 19. This supports investigating review staffing and document collection. One week's waiting files cannot establish the dominant annual cause or the financial return from either change. If avoided hiring is unlikely, Bellhaven could defer paid AI expansion and prioritize document collection or a reviewer, while preserving approved information handling.

REVIEW ASSISTANCE COULD CHANGE THE TAX CHOICE

Testing extraction alone, review assistance alone and their combination can reveal whether AI reduces the senior burden. No price or measured saving is available for review assistance, so the audit assigns no additional profit to that opportunity.

TRUST ASSESSMENT / RECURRING PROFIT AND SHARED ASSUMPTIONS

# The Profit Calculations Reconcile but Remain Conditional

Bellhaven's supplied revenue and costs reconcile to \$9.6 million and \$8.3 million, leaving \$1.3 million operating profit. The annual gains below use mature operation at 200 bookkeeping clients, unchanged prices and acceptable quality. They exclude one-time costs and assume no valuable senior work is displaced.

| ANNUAL OPERATING ASSUMPTIONS                                  | RECOMPUTED GAIN | WHAT THE RESULT ESTABLISHES                                                               |
|---------------------------------------------------------------|-----------------|-------------------------------------------------------------------------------------------|
| 10% bookkeeping improvement; 64 staff licenses.               | \$11,120        | One avoided \$98,000 position less \$60,000 bookkeeping fees and \$26,880 staff licenses. |
| 20% bookkeeping improvement; 64 staff licenses.               | \$109,120       | Two avoided positions less the same tool and license charges.                             |
| 20% bookkeeping plus proportional tax savings.                | \$153,020       | Adds \$43,900 after the recurring tax fee.                                                |
| 20% bookkeeping plus contractor-first tax savings.            | \$193,620       | Adds \$84,500 after the recurring tax fee.                                                |
| The favorable tax allocation and no extra office license fee. | \$220,500       | Leaves \$29,500 below the \$250,000 goal before unpriced continuing costs.                |

## The Favorable Scenario Is Not a Fully Costed Forecast

The calculations deduct specified licenses and tax fees, but no quantified allowance establishes continuing integration, support, monitoring or additional review-assistance costs. Existing staff time is an economic burden when it displaces useful work; it is not automatically an extra cash expense. Any new recurring spending or lost contribution reduces the modeled return. The analysis correctly identifies this work but does not fully price it.

## Demand and Senior Workload Can Weaken the Return Together

With eight new bookkeeping clients and one genuinely avoided position, the contractor-first tax scenario with 64 licenses gives \$99,220. If all 650 verification hours also displace \$175 of contribution and bookkeeping prices fall 1% across 188 clients, the combined annual result is a loss of about \$46,904 before unpriced costs. This combined sensitivity is not a forecast or a probability estimate.

**A MIXED PROGRAM DOES NOT ESTABLISH AI ATTRIBUTION**

Adding the reviewer's \$53,450 to the favorable AI scenario produces \$273,950 only if demand, capacity and costs do not overlap. The reviewer's independent contribution cannot be relabeled as AI-generated profit. Any interaction requires a comparison that measures the combined program and its components.

TRUST ASSESSMENT / FIRST YEAR CASH

# The Cash Totals Fit Only Their Stated Deployment Schedule

The base schedule totals \$95,080: \$26,880 for staff licenses, \$45,000 for phased bookkeeping use, \$8,000 for setup and training, \$5,200 for historical tax testing and \$10,000 for legal review. Recalculation confirms the total and the \$24,920 remaining against the cash ceiling.

| FIRST-YEAR ARRANGEMENT                                  | RECOMPUTED CASH | MEANING AGAINST THE \$120,000 CEILING                                                                                  |
|---------------------------------------------------------|-----------------|------------------------------------------------------------------------------------------------------------------------|
| The base program with 30 staff licenses.                | \$80,800        | \$39,200 remains. Suitability for the intended staff and tasks still needs confirmation.                               |
| The base program with approved included office access.  | \$68,200        | \$51,800 remains. The separate bookkeeping tool is still paid.                                                         |
| The 64-license base plus full live tax-processing fees. | \$110,680       | \$9,320 remains before new-client charges and other additions.                                                         |
| Included office access plus the seasonal reviewer.      | \$117,600       | \$2,400 remains. This arrangement includes historical tax testing, not full live tax processing.                       |
| Included office access plus reviewer and live tax fees. | \$133,200       | Exceeds the ceiling by \$13,200 before further additions.                                                              |
| Stop after stage two with 64 annual staff licenses.     | \$63,580        | Removes \$31,500 of later bookkeeping charges; this is a scheduled-cost comparison, not a verified cancellation quote. |

## New Clients Can Consume the Reviewer Package Budget Margin

At \$25 per additional client-month, the \$2,400 margin covers only 96 additional client-months. If all 20 new clients use the tool throughout the seven later months, charges rise by \$3,500 and the reviewer package reaches \$121,100, or \$1,100 above the ceiling. Arrival dates therefore change affordability.

## Cash Timing Must Match the Work That Generates the Benefits

The bookkeeping charges cover 11 operating months: two at 90 clients, two at 180 and seven at 180. A dated plan must specify the preparation period and confirm whether any charge applies during the remaining month. Mature profit assumes 200 clients, so neither first-year profit nor achievement by month 24 follows from the recurring table alone.

THE REVIEWER AND FULL TAX ROLLOUT NEED SEPARATE FUNDING DECISIONS

The \$117,600 package cannot support an assumption that full live tax processing is already funded. Stage commitments around actual client arrivals, tax-season restrictions and quotations. A nominal budget margin does not establish that all planned work is affordable.

TRUST ASSESSMENT / ATTRIBUTION AND IMPLEMENTATION

# A Useful Pilot Requires a Defined Comparison and Staff Capacity

The general evaluation method is sound: measure acceptable completed work, total labor and financial changes against what would happen without AI. The official AI evaluation guidance also separates technical performance from business impact. Bellhaven's proposed pilot still needs several decisions before it can support a reliable financial conclusion.

| IMPLEMENTATION QUESTION                           | WHAT THE AUDIT FINDS                                                                                                                  | CONSEQUENCE FOR THE PILOT                                                                                                                                                   |
|---------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| What does the comparison isolate?                 | A firm-wide approved AI service may also affect the teams used as the bookkeeping comparison.                                         | Record access and usage in both groups. A shared staff service makes the contrast an estimate of added bookkeeping-tool value, not necessarily all AI value.                |
| How long does an untreated comparison remain?     | The schedule moves from two participating teams to all four after two months.                                                         | Retain a defensible comparison through the decision period or state the weaker basis for later estimates. Continued measurement alone does not preserve causal attribution. |
| Can the first review establish the annual return? | The first 45-day decision precedes completion of the four-month pilot sequence.                                                       | Use the first review to stop, repair or continue a limited test. Annual results require later observation and reconciliation with staffing and accounts.                    |
| Can management deliver the planned work?          | Partner time is capped at six hours weekly. A proposed allocation exists; routine operational effort is uncoded.                      | Name accepted owners, their available hours and displaced duties before authorizing concurrent bookkeeping, tax and review tests.                                           |
| Have safeguards been demonstrated?                | The plan specifies restricted access, qualified review, recorded actions and manual recovery. No deployed configuration was examined. | Historical files still need approved handling. Require actual access, recovery and quality tests before live client work.                                                   |

## The Adoption Plan Addresses Distinct Staff Barriers

The survey has 58 responses from 64 staff, with 19 reporting weekly use. Among the other 39 respondents, 22 report unclear permission, 14 lack of know-how and nine low usefulness; those overlapping answers cannot be added as separate people. Four reports of client information in personal tools justify prompt investigation, but do not establish the full extent of exposure.

The proposed response appropriately combines permitted access, task-specific practice, quality assessment and assignment of saved hours. No layoffs for 12 months, credentialed review and the January 15 to April 15 restriction on new live tax workflows remain binding requirements.

TRUST ASSESSMENT / REVIEW FINDINGS AND NEXT DECISIONS

# The Audit Supports Preparation and Limited Tests

This examination recomputed the principal financial and pilot calculations, compared the recommendations with the stated goals and tested competing explanations and combined adverse conditions. The principal arithmetic reconciles. The consequential gaps concern full costs, staffing and demand, cash timing, comparison design and demonstrated operating controls.

## The Review Confirms Reasoning Within a Defined Scope

The operating figures are supplied inputs, not independently verified company records. Selected accounting and workplace research, the developer-study update, official evaluation guidance, National Institute of Standards and Technology risk guidance and Federal Trade Commission safeguards guidance were examined directly. The accounting research supports the stated distinction between association and experiment. The selected guidance supports evaluation and control principles; it does not certify Bellhaven's implementation.

The remaining research references and all current product, fee-guidance and jurisdiction-specific assertions were not fully reverified. No live pilot, supplier configuration, original company ledger or separate reviewer's complete work was available. This is a fresh analytical examination of shared inputs, not independent verification of business outcomes or certification of earlier review activity.

| ACTION STATUS                  | SPECIFIC CONSEQUENCE AND CONDITIONS                                                                                                                                                                                                                            |
|--------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Justified now.                 | Prepare the restricted test, confirm the hiring plan and signed demand, investigate personal-tool use, and obtain approved-service and reviewer terms. These steps resolve questions that can change the next commitment.                                      |
| Suitable for a limited trial.  | Run approved bookkeeping comparisons and historical tax tests with named owners, accepted quality criteria, available staff hours, a spending ceiling and a tested fallback. A critical confidentiality or unauthorized-action failure stops the affected use. |
| Requires further confirmation. | Defer wider deployment, automatic posting and commitments relying on two avoided positions or the annual profit target until observed quality, usable capacity, demand and a fully reconciled budget support them.                                             |

## The Next Review Should Resolve the Commitment Questions

Proposed finance and operations owners should confirm demand, staffing without AI, client arrival dates and the dated cost plan during Days 1 to 10. The proposed operational and information-security owners should establish comparison rules, permitted data, quality standards and fallback before live testing. Management should decide by Day 45 whether to stop, repair or continue; unresolved material conditions justify narrower testing or deferral.

## Two Focused Reviews Would Provide Additional Assurance

A separate finance review using actual hiring plans, quotations and monthly workloads should test whether the savings and cash commitments can coexist. An observed controlled trial should test whether acceptable work improves enough to change spending after review, exceptions and recovery are included. Neither review has been completed.

WHAT THE READER SHOULD TAKE AWAY

The analysis is useful for choosing and structuring AI tests, and its principal calculations are dependable within their stated assumptions. Bellhaven has credible opportunities, but achievable gains and total implementation costs remain uncertain. Proceed with preparation and limited approved tests; let measured quality, usable capacity and verified finances determine expansion.