Akholi

PART 5 / 2026 RESEARCH STUDY

Manual Alteration by the Delivery Team

Agentic AI Outsourcing at Tier 1 Financial Institutions

Phil Hatch

150 engagements · 850 delivery-team respondents

Observational, cross-sectional study of a nonrepresentative sample. Financial institutions did not independently validate the reported experience. The findings do not establish causation or legal breach.

Version 1.8.2 | November 1, 2026

REPORT INFORMATION

Agentic AI Outsourcing at Tier 1 Financial Institutions

Full Study Report

Author	Phil Hatch
Publication date	November 1, 2026
Version	1.8.2
ISBN	979-8-234-26882-2
SSRN DOI	https://doi.org/10.2139/ssrn.7164559
Figshare DOI	https://doi.org/10.6084/m9.figshare.34036890
Contact	info@akholi.com
Website	akholi.com

AUTHOR ONLINE

Akholi author page | ORCID | LinkedIn | ResearchGate

SUGGESTED CITATION

Hatch, Phil. 2026. *Agentic AI Outsourcing at Tier 1 Financial Institutions: Full Study Report*. Version 1.8.2. Akholi. https://doi.org/10.2139/ssrn.7164559.

COPYRIGHT

© 2026 Phil Hatch and Akholi. All rights reserved.

No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the copyright holders, except for brief quotations with full attribution in reviews, research, and other noncommercial uses permitted by copyright law. Requests for permission should be sent to info@akholi.com.

USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence tools were used to prepare this report. Appendix G discloses how they were used.

Contents

Manual Alteration by the Delivery Team **5**
The Five Manual Alteration Methods **9**
1 Manual Alteration of Financial-Institution-Provided Data 11
2 Manual Alteration of Financial Institution Processes 17
3 Manual Alteration of AI Outputs 23
4 Reporting Alteration 29
5 Third-Party Work Done by Hand 39

Manual Alteration by Cancellation Status and Engagement Indicator Mean **43**
Distribution of Manual Alteration Across Roles **45**
1 Engagement Managers and Portfolio Directors 49

Reasons Given for Manual Alteration **51**
1 Reasons for Manual Alteration of Financial-Institution-Provided Data 55
2 Reasons for Manual Alteration of Financial Institution Processes 57
3 Reasons for Manual Alteration of AI Outputs 59
4 Reasons for Reporting Alteration 61
5 Reasons for Third-Party Work Done by Hand 63

Disclosure of Manual Alteration **65**
1 Disclosure to Vendor Management 69
2 Disclosure to the Financial Institution 77
3 Sources of the Disclosures That Occurred 83

Recording of Manual Alteration **85**
1 Systems of Record 87
2 Visibility in Reported Performance 89

Regulatory and Legal Context Reported by Respondents **91**
1 Regulatory Items Named by Respondents 93
2 Relationship to the Work Subject to Manual Alteration 95

Regulatory and Legal Knowledge of the Delivery Team **97**
1 Regulatory Understanding of the Respondent and Management 99
2 Financial Institution Discussion of Obligations With the Delivery Team 105
3 Regulatory Training 109

About the Author **113**
About Akholi **115**

“

This is the one area I would hold up as an example — the monitoring we added after last year's provider upgrade scare has held, and the model vendor has not surprised us once. What has not improved is our own memory of the manual work. The people who used to fix awkward LP questionnaires by hand have moved on, and I am not certain we could go back to that if we had to.”

PORTFOLIO DIRECTOR

CXM, Corporate and back-office functions

Manual Alteration by the Delivery Team

Respondents on all 150 engagements reported manual alteration as part of delivery. Their accounts of limited disclosure and incomplete recording raised concerns about how fully reported performance represented the work being done. Manual alteration, the study's term for the delivery team's changes to data, processes, outputs, or reporting and its performance of third-party work by hand, often involved several methods on the same engagement. Respondents described corrections that sustained delivery but were not always visible in engagement records or reports to the financial institution. Their unfavorable assessments of regulatory training added a concern about preparation for work subject to legal and regulatory requirements. Together, these findings describe a gap between the work respondents reported performing and the visibility and preparation surrounding that work.

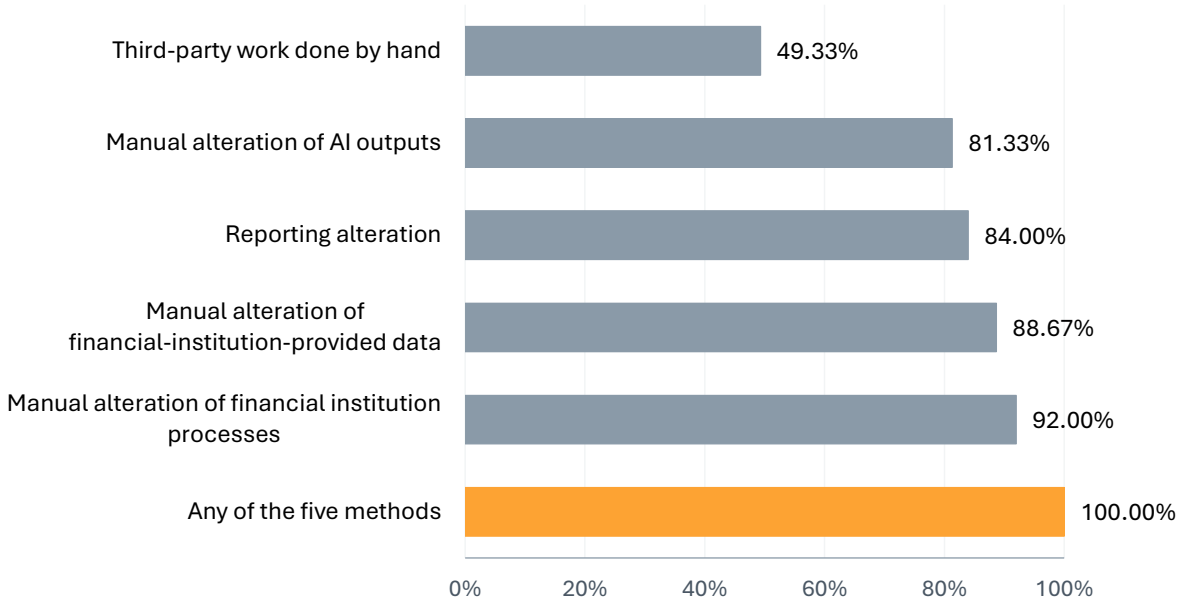
The five manual alteration methods are manual alteration of financial-institution-provided data, manual alteration of financial institution processes, manual alteration of AI outputs, reporting alteration, and third-party work done by hand. They are distinct practices, may overlap, and were measured differently. The occurrence counts combine direct True answers with written descriptions of a method on the respondent's engagement. Related ratings describe frequency, capability, knowledge, or controls rather than independently establishing occurrence.

The study collected the accounts of 850 members of vendor delivery teams on 150 engagements between January and June 2026. Its cross-sectional design recorded each engagement once. The results describe this study set and do not establish industry prevalence. The accounts do not independently establish what the financial institution knew or whether a legal or contractual obligation was breached.

Figure 5.1. Engagements with each manual alteration method, 150 engagements

100%

150 of 150 engagements had reported manual alteration.



Note: An engagement is counted for a method when at least one respondent answered True to its yes-or-no question or described that method in a written answer about the engagement. Manual alteration of data, processes, and AI outputs uses both sources. Reporting alteration and third-party work done by hand use written answers. An engagement may be counted under several methods. The base is 150 engagements.

Respondents reported manual alteration of financial institution processes on 138 of the 150 engagements (92.00%) and manual alteration of financial-institution-provided data on 133 (88.67%). They reported reporting alteration on 126 (84.00%), manual alteration of AI outputs on 122 (81.33%), and third-party work done by hand on 74 (49.33%).

Several methods were usually reported on the same engagement. Respondents described all five on 51 engagements (34.00%) and four on 65 (43.33%). They reported a single method on eight engagements.

Across the structured and written answers, 738 of the 850 respondents reported at least one method (86.82%). The method-based compilation included 2,410 written answers from 725 respondents on all 150 engagements. In 475 answers from respondents on 131 engagements, manual alteration was described as occurring every cycle or on most days. These are overlapping answer counts, not counts of separate incidents.

> *i do this almost every day now. i inform my team lead please, not the client, therefore it stays internal. it is exhausting but i see no other way."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Corporate and back-office functions

These findings describe the delivery team's reported experience. They do not establish what the financial institution knew or whether a reported practice breached a contractual or regulatory obligation. Manual correction may have been approved or necessary to maintain delivery. Neither its occurrence nor its absence from a report independently establishes misconduct.

> *The vendor's dashboard reported average turnaround time per valuation without disclosing that a meaningful share of deliverables were quietly redone by senior staff after bankers rejected the initial agent output, understating real cycle time."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Corporate and back-office functions

“

Weekly status decks reported case-processing speed against targets set for the manual process, so faster agent throughput looked like a win even on cases the agent had gotten factually wrong and compliance later had to unwind.”

INDIVIDUAL CONTRIBUTOR

BPO, Financial crime and client onboarding

The Five Manual Alteration Methods

Manual alteration of processes and data was reported on more engagements than any other method, although the five methods commonly overlapped. The evidence also differed by method. Personal manual alteration of data, processes, and AI outputs was asked through yes-or-no questions, with separate ratings of regular manual alteration by the wider delivery team. Reporting alteration and third-party work done by hand were identified in written answers. Their related ratings assess respondents' judgments about reporting and the retention or replacement of manual practices. Table 5.1 distinguishes the occurrence counts from these related measures.

Table 5.1. The five manual alteration methods and their evidence

Method	Occurrence evidence and related ratings	Respondents	Engagements
Manual alteration of financial-institution-provided data	Occurrence: C199 True answers and written answers. Related rating: C203.	452	133
Manual alteration of financial institution processes	Occurrence: C205 True answers and written answers. Related rating: C209.	546	138
Manual alteration of AI outputs	Occurrence: C211 True answers and written answers. Related rating: C215.	328	122
Reporting alteration	Occurrence: written answers, including C055. Related ratings: C052–C054.	477	126
Third-party work done by hand	Occurrence: written answers. Related ratings: C190–C192.	108	74

Note: Counts identify respondents and engagements with at least one direct True answer or written description of the method on the engagement, from 850 respondents and 150 engagements. Counts overlap across methods. Related ratings do not add to the occurrence counts. C numbers identify survey questions, including rated indicators. The outsourcing models shown in the paper are business process outsourcing (BPO), information technology outsourcing (ITO), customer experience management (CXM), and knowledge process outsourcing (KPO).

1 Manual Alteration of Financial-Institution-Provided Data

Manual alteration of financial-institution-provided data was reported on 133 of the 150 engagements (88.67%), often as recurring correction rather than an isolated intervention.

For personal manual alteration, the direct yes-or-no question recorded 421 True answers among 839 valid answers (50.18%), representing 127 engagements. The broader method-based compilation across written questions included 588 answers from 382 respondents on 128 engagements. Of those answers, 573 had negative sentiment and 223 described manual alteration every cycle or on most days.

Sentiment classifies an answer's reported effect on the engagement or respondent on five values: −10, −5, 0, +5, and +10. Negative values indicate harm, zero a neutral assessment, and positive values benefit. An answer may have an overall sentiment score and separate scores for the survey subjects it addresses. These classifications describe the account and do not establish its truth, the legality of a practice, or the adequacy of disclosure.

> *"I fix wrong Data every day, team lead know but we do not raise it with client, feedbacks would go nowhere anyway"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

> *"I flag it every week, my manager knows but nothing changes upstream. This Vendor takes zero ownership so we are the ones patching data every single day"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Technology engineering and operations

> *"i correct the extracts before loading them, half the fields come swapped or blank before we even open the file,"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Technology engineering and operations

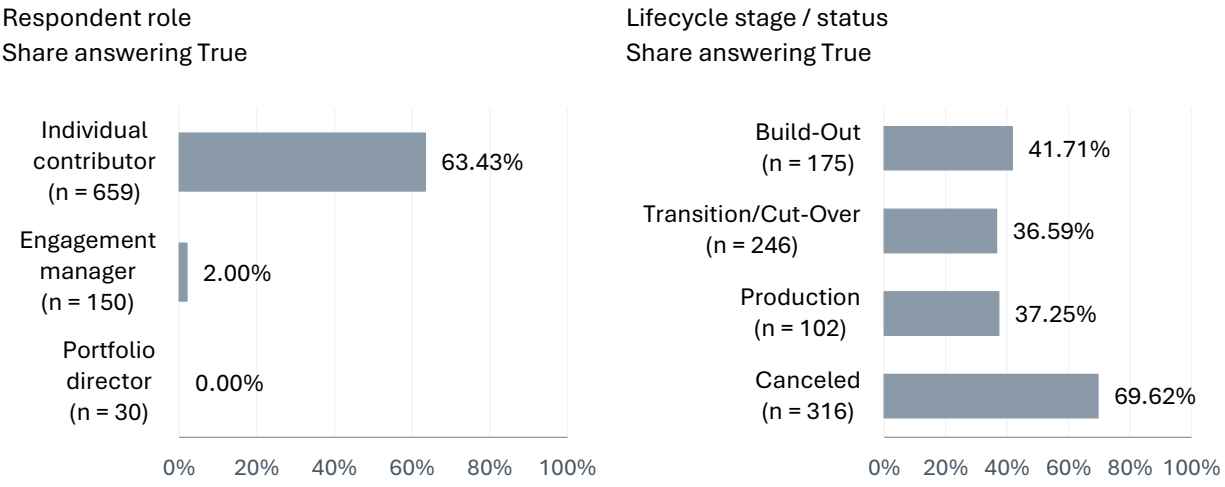
1.1 Personal Manual Alteration of Financial-Institution-Provided Data

Indicator C199 · I have altered client-provided data · Full statistical analysis ›

Half of the respondents answering the personal data question reported that they had manually altered financial-institution-provided data: 421 of 839 answered True (50.18%), while 418 answered False (49.82%). The True share was 69.62% among respondents on canceled engagements (n = 316 respondents), compared with 36.59% among those on active engagements in Transition/Cut-Over (n = 246). The notation n identifies the number of respondents or engagements in an analysis. These respondent-level shares describe personal answers, not the proportion of engagements assigned to a comparison group.

Figure 5.2. Share answering True to "I have altered client-provided data", by respondent role and engagement status

Personal manual alteration of financial-institution-provided data differed by role



Note: True or Yes answers: 421 of 839 answers (50.18%).
n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

The four core measures are cancellation rate, customer satisfaction, engagement indicator mean, and intent to leave. Cancellation rate is the share of a stated group of engagements canceled by the study date. Customer satisfaction is the figure recorded for each engagement on a 100-point scale, with the comparisons here restricted to active engagements. Engagement indicator mean summarizes respondents' ratings across the 100 engagement indicators, averaged within each engagement. The mean is the arithmetic average of the recorded values, and the median is the middle value in an ordered distribution. Intent to leave is the respondent's rating of the likelihood of leaving the current job within 12 months.

For categorical comparisons, an engagement is assigned the most frequent answer among its respondents, termed the plurality answer. Sentinel categories such as a preference not to answer remain answer categories under the existing rule, and tied engagements are excluded. This assignment differs from counting an engagement whenever any respondent reports a manual alteration method. An engagement-level analysis uses one summarized observation per engagement, while a respondent-level analysis uses paired answers from individuals.

A chi-square test examines whether a categorical distribution differs from independence. Welch F compares group means while allowing their variances, or spread of ratings, to differ. Degrees of freedom, df, describe the information available to estimate variation. A p value is the probability, under the null model, of a result at least as extreme as the observed result. Statistical significance does not establish the size, operational importance, or cause of a relationship.

Engagements assigned True on the personal data question had a higher observed cancellation rate and less favorable results on the other three core study measures. Cancellation was recorded on 61.84% of the 76 engagements assigned True, compared with 10.17% of the 59 assigned False (chi-square 37.19, df = 1, $p < 0.001$, n = 135 engagements). Among active engagements, customer satisfaction differed between the groups (Welch F 17.74, $p < 0.001$, n = 82 engagements). The groups also differed in engagement indicator mean (Welch F 93.60, $p < 0.001$) and intent to leave (Welch F 55.34, $p < 0.001$), each on 135 engagements. These associations describe reported experience measured after cancellation where an engagement had already been canceled. They do not establish whether manual alteration preceded cancellation or caused it.

Robustness checks repeat an analysis using alternative assumptions or selections. Their interpretation depends on the particular check and analytical unit. The linked indicator analyses provide the existing check results. The qualifications reported below, including failed or uncomputable checks and differences between engagement-level and respondent-level results, limit any general claim of stability.

The separate indicator-specific compilation for C199 contained 578 written answers from 378 respondents on 128 engagements, of which 571 received negative sentiment scores for this subject. Respondents described data arriving unusable despite earlier cleaning and access restrictions that prevented correction at the source. This indicator-specific count differs from the broader method-based compilation above and is not an additional set of incidents.

> *“Last cycle their file was missing tax ids for half the batch, I fixed it myself before run started… nothing gets checked before it lands on our desk, this happens basically every cycle now”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Corporate and back-office functions

> *Last cycle it was three whole cohort missing from one file, I only catch it because I count row against last cycle total and numbers didnt match up, if I hadnt counted it would have gone straight through with people missing pay.”*

INDIVIDUAL CONTRIBUTOR

BPO, Corporate and back-office functions

> *last month the cash statement came after cut off again… so i take previous day file and roll forward, thus the break clears on time.”*

INDIVIDUAL CONTRIBUTOR

BPO, Trade, settlement and reconciliation

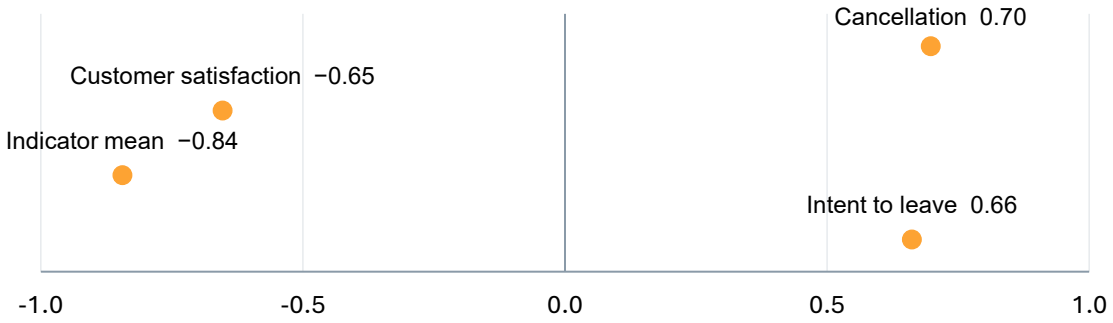
1.2 Regular Manual Alteration of Financial-Institution-Provided Data by the Wider Delivery Team

Indicator C203 · My greater team regularly alters client provided data · Full statistical analysis ›

Respondents described data manual alteration as a wider delivery-team practice as well as a personal activity. Agreement with the statement about regular manual alteration had a mean rating of 5.4 and a median of 5.0 (n = 773 respondents). The mean was 6.9 among respondents on canceled engagements (n = 287). Higher ratings indicate stronger agreement that regular manual alteration occurred. The question has no study-defined favorable or unfavorable direction, so these ratings are not classified as distressed.

Figure 5.3. Regular Manual Alteration of Financial-Institution-Provided Data by the Wider Delivery Team: associations with the study measures

Regular manual alteration of financial-institution-provided data by the wider delivery team: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = 0.70 150 engagements p < 0.001
Customer satisfaction	r = −0.65 94 active engagements p < 0.001
Indicator mean	r = −0.84 150 engagements p < 0.001
Intent to leave	r = 0.66 722 respondents p < 0.001

Note: Mean rating: 5.4 (n = 773 respondents).

Survey question: “My greater team regularly alters client provided data”.

Group means and numeric-answer bases: Individual contributor (n = 622): 5.70. Engagement manager (n = 131): 4.28. Portfolio director (n = 20): 2.45. Build-Out (n = 161): 4.68. Transition/Cut-Over (n = 232): 4.48. Production (n = 93): 4.24. Canceled (n = 287): 6.87.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational. A confidence interval expresses uncertainty under the stated method. A p value is the probability, under the null model, of a result at least as extreme as the observed result.

A mean is the arithmetic average and a median is the middle ordered rating. The coefficient r measures the direction and strength of a linear association, ranging from −1 to +1. A coefficient near either end describes a stronger linear relationship, while its sign identifies the direction. It does not establish causation.

The correlation coefficient r describes the direction and strength of a linear association. Stronger agreement was associated with higher cancellation rates (r = 0.70, n = 150 engagements) and a lower engagement indicator mean (r = −0.84, n = 150). Among active engagements, it was associated with lower customer satisfaction (r = −0.65, n = 94 engagements). At respondent level, stronger agreement was associated with higher intent to leave (r = 0.66, n = 722 respondents). The distinction between personal correction and a regular delivery-team practice is substantive, although these associations do not establish the sequence of events.

Of the 84 written answers coded to regular data manual alteration, 83 had negative sentiment. Respondents described data that required cleaning again after each delivery, with manual alteration known to the team lead but seldom logged or disclosed to the financial institution. The single positive answer described a data-exception procedure agreed with the financial institution's platform owner before correction.

> *"I keep note of fixes i make each sprint, my lead know about it, client does not. Take maybe extra hour every day and i do not think should be my job to clean their data before agent see it."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Technology engineering and operations

2 Manual Alteration of Financial Institution Processes

Manual alteration of financial institution processes was the most widely reported method, occurring in respondents' accounts on 138 of the 150 engagements (92.00%).

For personal manual alteration, the direct question recorded 490 True answers among 841 valid answers (58.26%), representing 132 engagements. The method-based compilation contained 857 written answers from 515 respondents on 137 engagements, including 826 classified as negative. In 143 answers, respondents described manual alteration every cycle or on most days.

> *"nobody upstream will say yes or no so i just go with my own call, decision take too long to come back and Payroll deadline is same every cycle so i cannot wait"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Corporate and back-office functions

> *"This is basically every week now, my manager knows about it. Ticket takes days to clear, fix itself takes an hour, so waiting on the ticket process makes no sense when the job needs doing"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Technology engineering and operations

2.1 Personal Manual Alteration of Financial Institution Processes

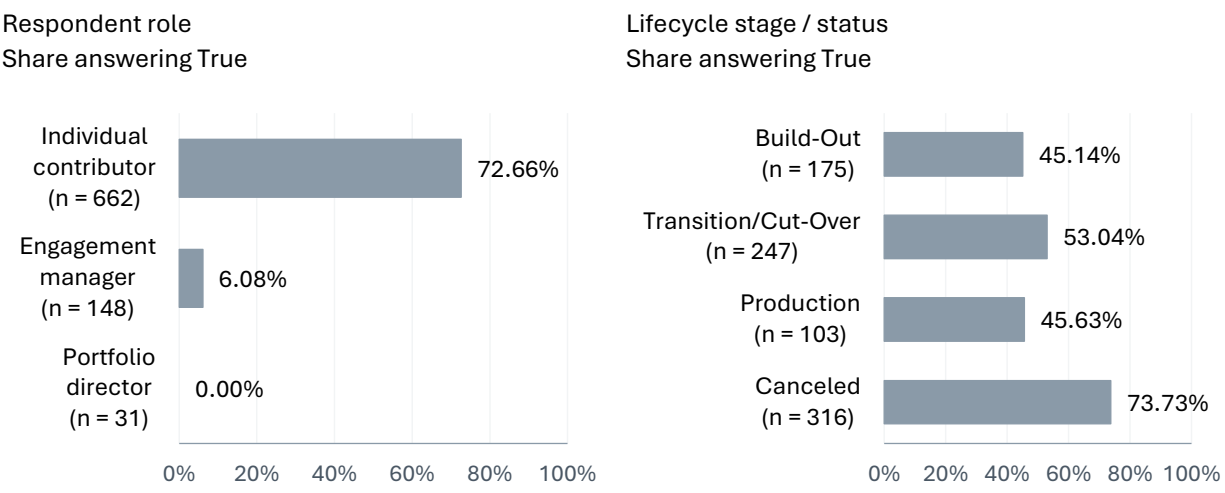
Indicator C205 · I have altered client processes · Full statistical analysis ›

A majority of respondents answering the personal process question reported manual alteration: 490 of 841 answered True (58.26%), while 351 answered False (41.74%). The True share was 73.73% among respondents on canceled engagements (n = 316), compared with 45.14% among those on active engagements in Build-Out (n = 175). By outsourcing model, it ranged from 69.42% among respondents on BPO engagements (n = 206) to 50.87% on CXM engagements (n = 230).

Figure 5.4. Share answering True to "I have altered client processes", by respondent role and engagement status

58.26%

490 of 841 answers reported personal process manual alteration. The remaining 351 answered False (41.74%).



Note: True or Yes answers: 490 of 841 answers (58.26%).
n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

Engagements assigned True on the process question had a cancellation rate of 50.00%, compared with 15.22% for engagements assigned False (94 and 46 engagements, respectively). The difference was supported by the existing categorical comparison (chi-square 15.77, df = 1, $p < 0.001$, n = 140 engagements). Active-engagement customer satisfaction also differed (Welch F 17.72, $p < 0.001$, n = 86 engagements), as did engagement indicator mean (Welch F 64.36, $p < 0.001$) and intent to leave (Welch F 36.02, $p < 0.001$), each on 140 engagements. These are associations with the assigned answer, not evidence that process manual alteration caused the outcomes.

The indicator-specific compilation for C205 contained 849 written answers from 513 respondents on 137 engagements, including 841 with negative sentiment scores for this subject. Respondents described delayed financial institution decisions, change approval designed for infrequent changes, and procedures that were undocumented or inconsistent with practice. They reported changing processes before recording the changes, or without recording them. The method-based compilation above and this indicator-specific compilation are separate summaries. Their different sentiment counts have not been reconciled and must not be interpreted as a whole and its subset.

> *“Escalating a break means waiting on someone who never actually rules on it, we just pick a call and run with it. And whatever that forces into the workflow gets written up later, if it gets written up at all.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

> *“i changed a step on the reconciliation process last week, nothing about the change process here really matches how the agent behaves day to day, so we adjust as we go and i discuss about it with my lead, its not a big drama”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

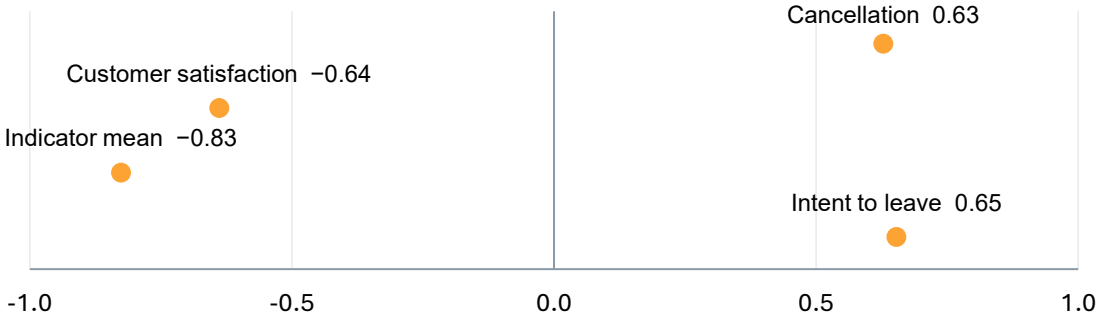
2.2 Regular Manual Alteration of Financial Institution Processes by the Wider Delivery Team

Indicator C209 · My greater team regularly alters client processes · Full statistical analysis ›

Regular process manual alteration received stronger agreement than the corresponding questions on data or AI outputs. Its mean rating was 5.7 and its median was 6.0 (n = 779 respondents). Respondents on canceled engagements reported a mean of 7.1 (n = 291). Higher ratings indicate stronger agreement that the wider delivery team regularly performed manual alteration. The question has no study-defined favorable or unfavorable direction.

Figure 5.5. Regular Manual Alteration of Financial Institution Processes by the Wider Delivery Team: associations with the study measures

Regular manual alteration of financial institution processes by the wider delivery team: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = 0.63 150 engagements p < 0.001
Customer satisfaction	r = −0.64 94 active engagements p < 0.001
Indicator mean	r = −0.83 150 engagements p < 0.001
Intent to leave	r = 0.65 730 respondents p < 0.001

Note: Mean rating: 5.7 (n = 779 respondents).

Survey question: “My greater team regularly alters client processes”.

Group means and numeric-answer bases: Individual contributor (n = 631): 6.00. Engagement manager (n = 130): 4.56. Portfolio director (n = 18): 3.00. Build-Out (n = 161): 4.63. Transition/Cut-Over (n = 230): 4.84. Production (n = 97): 5.23. Canceled (n = 291): 7.10.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Stronger agreement was associated with a lower engagement indicator mean (r = −0.83, n = 150 engagements) and a higher cancellation rate (r = 0.63, n = 150). Among active engagements, it was associated with lower customer satisfaction (r = −0.64, n = 94 engagements). It was also associated with higher intent to leave at respondent level (r = 0.65, n = 730 respondents).

Of the 111 written answers coded to regular process manual alteration, 110 had negative sentiment. Respondents described routine changes made by the delivery team when financial institution decisions were delayed. The team lead was described as knowing about these changes, while disclosure to the financial institution was uncommon in the accounts.

> *“This is normal now, my team decide small things ourselves and tell manager after not before. I don't ask if client know, not my place. Happens most weeks, only way to hit the cycle.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Corporate and back-office functions

“

What would help is one person owning this end to end. Right now delivery lead, the AI team and quality all give different instruction and we follow whichever came last. Second, when the assist gives a wrong answer to the floor it must be somebody's job to fix the source, not mine to patch it in the ticket, we do that all week and nothing changes upstream”

INDIVIDUAL CONTRIBUTOR

CXM, Corporate and back-office functions

3 Manual Alteration of AI Outputs

Manual alteration of AI outputs was reported on 122 of the 150 engagements (81.33%), although fewer respondents personally reported this method than personal manual alteration of data or processes.

For personal manual alteration, the direct question recorded 278 True answers among 828 valid answers (33.57%), representing 108 engagements. The method-based compilation contained 429 written answers from 289 respondents on 121 engagements. Of those answers, 413 had negative sentiment and 95 described manual alteration every cycle or on most days.

Within the method-based compilation, 14 written answers on nine engagements described instructions to continue processing outputs without raising the problem. The reasons compilation later in the paper reports a different count for this theme. The two summaries are identified separately because their coverage has not been reconciled.

> *Nobody trusts the raw output so I rewrite it every time, that's the real Job now not what they think we do"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Corporate and back-office functions

> *the problem is nobody talk about how often team do this because everyone just fix quietly on their own, do understand I only track for myself, I do it almost every case I recieve"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Corporate and back-office functions

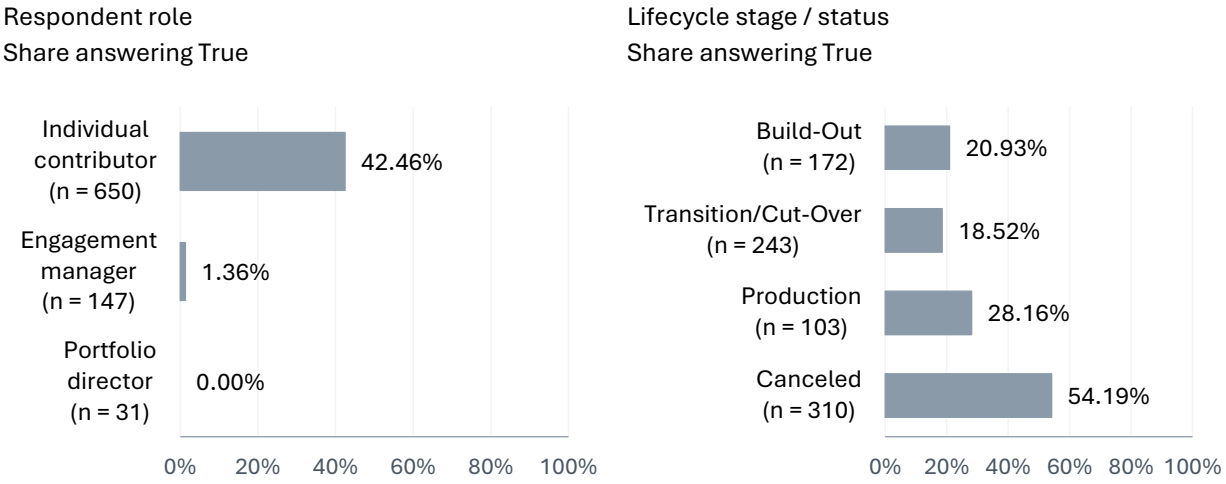
3.1 Personal Manual Alteration of AI Outputs

Indicator C211 · I have altered AI outputs · Full statistical analysis ›

One-third of respondents answering the personal AI-output question reported manual alteration: 278 of 828 answered True (33.57%), while 550 answered False (66.43%). The True share was 54.19% among respondents on canceled engagements (n = 310), compared with 18.52% among those on active engagements in Transition/Cut-Over (n = 243).

Figure 5.6. Share answering True to "I have altered AI outputs", by respondent role and engagement status

One-third of respondents answering the personal AI-output question reported manual alteration: 278 of 828 answered True (33.57%), while 550 answered False (66.43%).



Note: True or Yes answers: 278 of 828 answers (33.57%).
n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

The cancellation rate was 76.19% among the 42 engagements assigned True on the AI-output question, compared with 22.00% among the 100 assigned False (chi-square 36.86, df = 1, $p < 0.001$, n = 142 engagements). Active-engagement customer satisfaction differed between the groups (Welch F 11.44, p = 0.005, n = 88 engagements). Engagement indicator mean (Welch F 87.54, $p < 0.001$) and intent to leave (Welch F 79.99, $p < 0.001$) also differed, each on 142 engagements. The direction of these associations is consistent with the accounts of difficult delivery, but it does not establish the timing or causal role of manual alteration.

The indicator-specific compilation for C211 contained 407 written answers from 276 respondents on 116 engagements, including 399 with negative sentiment. Respondents described inconsistent outputs from repeated runs on the same input. They reported selecting and correcting outputs by hand before passing them onward because the approval process

was slower than the delivery deadline. These counts refer to C211 coding rather than the broader AI-output method compilation.

> *the problem is the ai gives a different match each time i run it, so i correct it myself before it moves on, on that basis i do not wait for approval"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

> *So what I do is run it, then run it again, and if the two agree I go with it, and if they don't I work it out by hand. That's my process. Nobody designed that, I invented it because I needed something."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

> *the problem is change control takes weeks for anything, so whats the point raising it every time, i patch the output myself and keep the queue moving, thats just how this job works and truthfully im quicker at catching it than the process is anyway"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Corporate and back-office functions

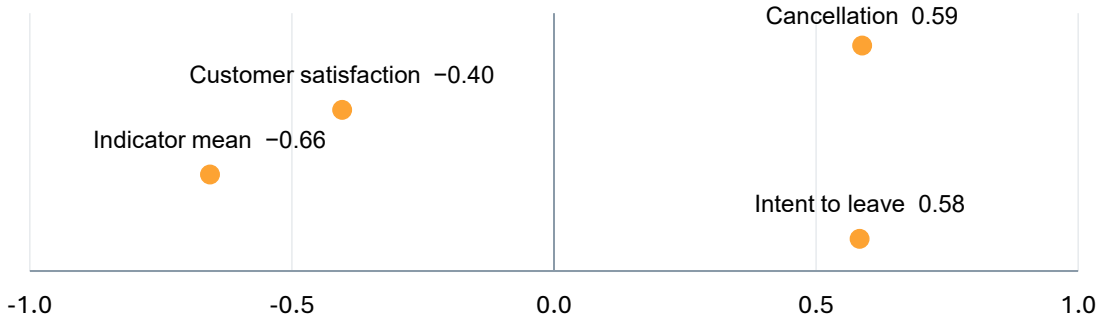
3.2 Regular Manual Alteration of AI Outputs by the Wider Delivery Team

Indicator C215 · My greater team regularly alters AI outputs · Full statistical analysis ›

Regular manual alteration of AI outputs received the weakest agreement among the three wider delivery-team questions. The mean rating was 4.4 and the median was 4.0 (n = 665 respondents), compared with means of 5.4 for data and 5.7 for processes. Respondents on canceled engagements reported a mean of 5.8 (n = 228). Higher ratings indicate stronger agreement that manual alteration was regular. The question has no study-defined favorable or unfavorable direction.

Figure 5.7. Regular Manual Alteration of AI Outputs by the Wider Delivery Team: associations with the study measures

Regular manual alteration of AI outputs by the wider delivery team: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = 0.59 150 engagements p < 0.001
Customer satisfaction	r = −0.40 94 active engagements p < 0.001
Indicator mean	r = −0.66 150 engagements p < 0.001
Intent to leave	r = 0.58 620 respondents p < 0.001

Note: Mean rating: 4.4 (n = 665 respondents).

Survey question: “My greater team regularly alters AI outputs”.

Group means and numeric-answer bases: Individual contributor (n = 556): 4.58. Engagement manager (n = 99): 3.59. Portfolio director (n = 10): 2.50. Build-Out (n = 150): 3.41. Transition/Cut-Over (n = 201): 3.61. Production (n = 86): 4.35. Canceled (n = 228): 5.77.

r is Pearson’s correlation coefficient, measuring signed linear association from –1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Stronger agreement was associated with a higher cancellation rate (r = 0.59, n = 150 engagements) and higher intent to leave (r = 0.58, n = 620 respondents). Its association with customer satisfaction among active engagements was negative (r = −0.40, n = 94 engagements) and weaker than for the corresponding data and process questions. Agreement was also associated with a lower engagement indicator mean (r = −0.66, n = 150 engagements).

All 23 written answers coded to this indicator had negative sentiment. The 19 that addressed AI-output manual alteration described continuing delivery-team correction before outputs were seen outside the delivery team, with little recording of that work.

> *“it's f***ing unpredictable, same input different answer. i fix it quiet every time, my lead knows, client doesn't. that's just how it is here.”*
>
> **INDIVIDUAL CONTRIBUTOR**
> CXM, Corporate and back-office functions

“

The vendor's monthly performance report calculates first-contact resolution using only tickets the agent both received and closed within the same reporting window, so any ticket the agent picks up late in the month and carries into the next cycle simply disappears from both months' denominators instead of counting as a miss.”

INDIVIDUAL CONTRIBUTOR

CXM, Technology engineering and operations

4 Reporting Alteration

Respondents described reporting alteration on 126 of the 150 engagements (84.00%), including methods that obscured manual correction or overstated agent-completed work.

In the reporting-method question, 484 of 775 respondents named a method (62.45%), while 291 answered “I do not know.” A separate compilation of descriptions concerning the respondent’s own engagement contained 486 written answers from 477 respondents on 126 engagements, including 485 with negative sentiment. Naming a method and describing it on one’s own engagement are different count populations.

> *“The vendor rolled unresolved reconciliation exceptions into the next period's dashboard so the error count never showed as owned, and it took a manual pull of the raw exception log to see the real backlog.”*
>
> **ENGAGEMENT MANAGER**
>
> BPO, Corporate and back-office functions

4.1 Reporting Methods Described by Respondents

Substituting activity counts for outcomes was the most frequently described reporting method, appearing in 112 answers on 78 engagements. This method is termed proxy substitution. Category conflation, which counted human-corrected work as the agent's, appeared in 52 answers on 45 engagements. Denominator exclusion, which removed failures from the base of a reported rate, appeared in 50 answers on 39 engagements. Table 5.2 defines the reporting methods identified in the coded written answers.

Table 5.2. Reporting methods described in the written answers

Reporting method	Description	Answers	Engagements
Proxy substitution	An activity count stands in for an outcome.	112	78
Category conflation	Human-corrected, sandbox or overridden work is counted as agent-completed production work.	52	45
Denominator exclusion	Failures or exceptions are removed from the base of a reported rate.	50	39
Scope restriction	The measure covers only the easy or parseable subset.	38	32
Manual effort omitted from reported measures	The labor does not appear in any reported measure.	27	25

Table 5.2. Reporting methods described in the written answers (continued)

Reporting method	Description	Answers	Engagements
Exceptions omitted from financial institution reports	The exception is recorded internally but omitted from what the financial institution receives.	23	20
Status misrepresentation	A readiness or completion status is reported falsely, letting a gate pass.	20	14
Validation-set selection	Accuracy is reported on a set chosen to exclude the hardest cases.	18	16
Clock reset	Elapsed-time measurement restarts on rework so delay never accrues.	4	3
Aging suppressed	Open items are pooled into a category that does not age, so duration never surfaces.	3	3

Note: The base is 486 written answers describing reporting alteration on the respondent's own engagement, covering 126 engagements. An answer may describe several methods. The method labels summarize coded descriptions, rather than independently verified reporting practices.

> *"The vendor rolled settlement fails that were manually corrected by a human into the same success bucket as fails the agent resolved on its own, which overstated how much the agent was actually doing correctly."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

> *The vendor's settlement dashboard reported the fails rate as a percentage of trades processed that day rather than trades still open, so a growing backlog of unresolved fails never showed up as a rising number.”*

ENGAGEMENT MANAGER

BPO, Trade, settlement and reconciliation

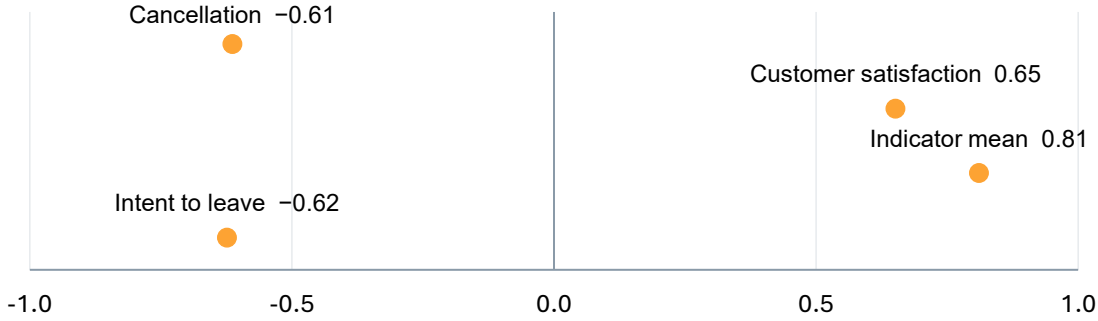
4.2 Accuracy of Dashboards and Performance Reports

Indicator C052 · The dashboard and performance reports given to the customer are accurate · Full statistical analysis ›

Respondents' ratings of dashboard and performance-report accuracy indicated substantial reservations about the reports the financial institution received. The mean rating was 4.8 and the median was 5.0 (n = 794 respondents). Ratings were distressed in 220 answers (27.71%). Respondents on canceled engagements reported a mean of 3.5 (n = 301). These ratings describe perceived accuracy rather than an independent audit of the reports.

Figure 5.8. Accuracy of Dashboards and Performance Reports: associations with the study measures

Accuracy of dashboards and performance reports: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.61 150 engagements p < 0.001
Customer satisfaction	r = 0.65 94 active engagements p < 0.001
Indicator mean	r = 0.81 150 engagements p < 0.001
Intent to leave	r = −0.62 745 respondents p < 0.001

Note: Survey question: "The dashboard and performance reports given to the customer are accurate".

Distressed ratings: 220 of 794 answers (27.71%).

Group means and numeric-answer bases: Individual contributor (n = 615): 4.66. Engagement manager (n = 150): 5.16. Portfolio director (n = 27): 6.33. Build-Out (n = 162): 5.85. Transition/Cut-Over (n = 233): 5.46. Production (n = 98): 5.67. Canceled (n = 301): 3.47.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Distressed ratings are ratings of 3 or below on indicators for which a higher rating is favorable. The percentage uses valid respondent ratings as its base. An engagement with a distressed mean is a separate engagement-level classification. For indicators converted to the favorable direction, the same threshold applies after conversion.

Higher perceived report accuracy was associated with a lower cancellation rate (r = −0.61, n = 150 engagements) and a higher engagement indicator mean (r = 0.81, n = 150). It was also associated with higher customer satisfaction among active engagements (r = 0.65, n = 94 engagements) and lower intent to leave (r = −0.62, n = 745 respondents). The engagement mean for this indicator was distressed on 29 of the 150 engagements.

The indicator-specific compilation contained 468 written answers from 453 respondents on 125 engagements, including 467 with negative sentiment. Respondents described measures that counted a task as complete when an agent acted, excluded cases requiring human correction, or blended and reclassified failures. The single positive answer concerned figures the respondent could substantiate.

> *"The vendor's status dashboard counted a task as complete the moment the agent submitted an action, before any human check confirmed the pay run was correct, so failed items looked closed until someone dug into the underlying case log."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Corporate and back-office functions

> *"The vendor's weekly report calculates average handle time using only tickets the agent fully resolves on its own, so the growing share of tickets the agent partially processes and then silently hands back to a human queue never enters the handle-time calculation at all, keeping the headline average artificially low."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Corporate and back-office functions

> *"The vendor's monthly billing accuracy dashboard reports fee exceptions only after they clear a manual override queue, so exceptions the agent catches and auto-corrects same-day never post as exceptions at all, making the vendor's own error rate look better than the fee leakage the agent is actually finding."*
>
> **ENGAGEMENT MANAGER**
>
> BPO, Payments, cards and disputes

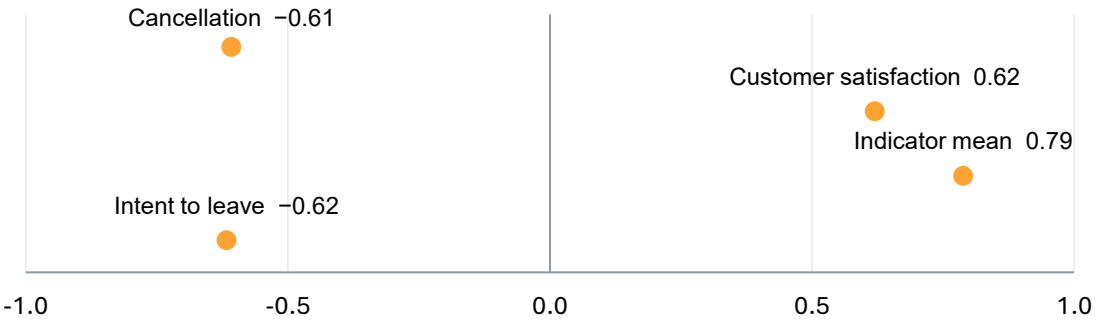
4.3 Reporting Alteration in Dashboards and Performance Reports

Indicator C053 · The vendor manipulates the dashboard and performance reports given to the customer · Full statistical analysis ›

Respondents' assessments of reporting alteration were less favorable on canceled engagements. After conversion to the favorable direction, the overall mean rating was 4.7 and the median was 5.0 (n = 786 respondents). The mean among respondents on canceled engagements was 3.4 (n = 302). The conversion subtracts the original rating from 11, so a higher converted rating indicates less perceived reporting alteration. It does not turn a reported judgment into a verified occurrence count.

Figure 5.9. Reporting Alteration in Dashboards and Performance Reports: associations with the study measures

Reporting alteration in dashboards and performance reports: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.61 150 engagements p < 0.001
Customer satisfaction	r = 0.62 94 active engagements p < 0.001
Indicator mean	r = 0.79 150 engagements p < 0.001
Intent to leave	r = −0.62 739 respondents p < 0.001

Note: Survey question: "The vendor manipulates the dashboard and performance reports given to the customer".

Distressed ratings: 235 of 786 answers (29.90%).

Group means and numeric-answer bases: Individual contributor (n = 613): 4.61. Engagement manager (n = 150): 5.19. Portfolio director (n = 22): 5.41. Build-Out (n = 168): 5.52. Transition/Cut-Over (n = 220): 5.56. Production (n = 96): 5.78. Canceled (n = 302): 3.38.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The indicator is converted to the favorable direction as 11 minus the original rating before these comparisons.

Higher converted ratings were associated with a lower cancellation rate (r = −0.61, n = 150 engagements) and a higher engagement indicator mean (r = 0.79, n = 150). They were also associated with higher active-engagement customer satisfaction (r = 0.62, n = 94 engagements) and lower intent to leave (r = −0.62, n = 739 respondents). Distressed ratings were recorded in 235 of the 786 answers (29.90%), and the engagement mean was distressed on 33 of the 150 engagements. For this converted indicator, distressed ratings of 3 or below correspond to original ratings above 7.

The indicator-specific compilation contained 465 written answers from 456 respondents on 125 engagements, including 464 with negative sentiment. Respondents described the use of activity counts, exclusion of failures, and attribution of human-corrected work to the agent in ways that maintained favorable reported figures.

> *"Purchase requisitions the agent could not match to a valid supplier record were logged as processed with a manual override flag that never surfaced in the weekly summary the customer received."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Corporate and back-office functions

> *"The vendor's exception dashboard groups unresolved custody breaks under a catch-all pending category that does not age, so a break sitting open for weeks looks the same as one opened yesterday."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

> *"The vendor's weekly ops dashboard tracks fee break resolution time starting from ticket creation rather than from when the agent first flagged the anomaly internally, so breaks the agent catches and holds for days before a ticket gets opened never count against the resolution clock at all."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Payments, cards and disputes

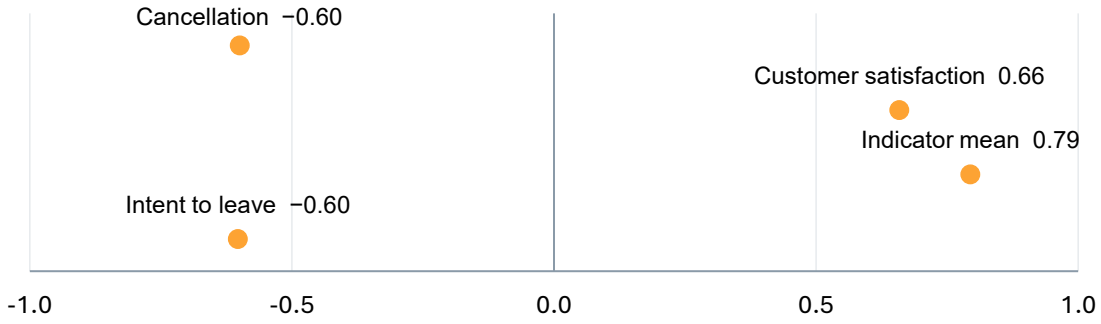
4.4 Retrospective Comparison of Reporting Alteration

Indicator C054 · Compared to 5 years ago, is manipulation of the dashboards and performance reports given to the customer more or less common now? · Full statistical analysis ›

Respondents' retrospective judgments of reporting alteration were less favorable on engagements with less favorable reported outcomes.

Figure 5.10. Retrospective Comparison of Reporting Alteration: associations with the study measures

Retrospective comparison of reporting alteration: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.60 150 engagements p < 0.001
Customer satisfaction	r = 0.66 94 active engagements p < 0.001
Indicator mean	r = 0.79 150 engagements p < 0.001
Intent to leave	r = −0.60 747 respondents p < 0.001

Note: Survey question: “Compared to 5 years ago, is manipulation of the dashboards and performance reports given to the customer more or less common now?”.

Distressed ratings: 247 of 795 answers (31.07%).

Group means and numeric-answer bases: Individual contributor (n = 619): 4.67. Engagement manager (n = 150): 5.25. Portfolio director (n = 24): 5.96. Build-Out (n = 164): 5.65. Transition/Cut-Over (n = 232): 5.57. Production (n = 98): 5.95. Canceled (n = 301): 3.43.

r is Pearson’s correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

The indicator is converted to the favorable direction as 11 minus the original rating before these comparisons.

The converted mean rating was 4.8 and the median was 5.0 (n = 795 respondents). A higher converted rating indicates less perceived reporting alteration relative to five years earlier. It was associated with a higher engagement indicator mean (r = 0.79, n = 150 engagements) and a lower cancellation rate (r = −0.60, n = 150 engagements). Among active engagements, it was associated with higher customer satisfaction (r = 0.66, n = 94 engagements). At respondent level, it was associated with lower intent to leave (r = −0.60, n = 747 respondents). Ratings were distressed in 247 answers (31.07%), and the engagement mean was distressed on 35 of the 150 engagements.

No written answer addressed this comparison. The figures retain the full answering base and do not establish that every respondent had sufficient work experience to make a five-year comparison. They describe retrospective judgments, not a measured change over five years. The present results must not be interpreted as an analysis restricted to respondents with an established five-year work history. Part 3, “Perceived Reporting Change Compared With Five Years Earlier,” section 2.3 under “Oversight of Reporting Provided to the Financial Institution,” uses an experience-eligible base of 176 respondents and examines the governance, oversight, and control score. That analysis has a different base and outcome from the full-answering-base comparisons reported here. Its raw direction also differs from the favorable-direction conversion used in this section.

5 Third-Party Work Done by Hand

Third-party work done by hand was described on 74 of the 150 engagements (49.33%), making it the least widely reported of the five methods.

The method-based compilation included 113 written answers from 108 respondents on those engagements, of which 95 answers had negative sentiment. Most descriptions appeared in the comment question on the third-party ecosystem (84 of 113 answers).

5.1 Third-Party Work Described by Respondents

Respondents described performing work by hand when third-party systems did not deliver it in a form an agent could use. They reported retrieving records individually from provider portals without an interface, entering data when feeds failed or changed format, and repeating work after unannounced provider changes. In 31 of the 113 answers, the work occurred every cycle or on most days. In 18 answers, respondents said the change became apparent only after a downstream failure.

> *“administrator portal is same as day one… still no proper interface, we pull screens by hand at cut off.”*
>
> **INDIVIDUAL CONTRIBUTOR**
> BPO, Trade, settlement and reconciliation

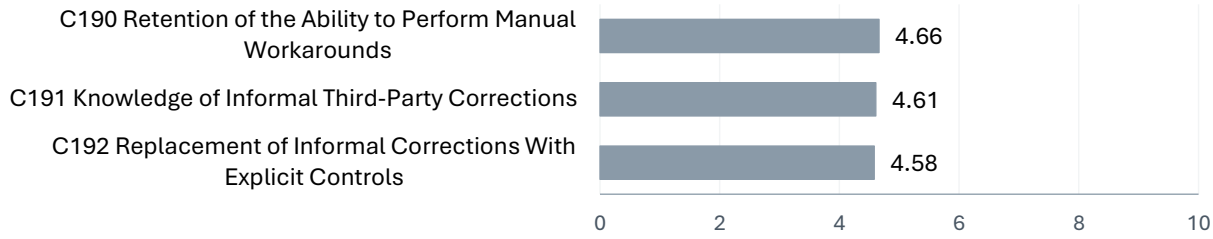
> *“Everything we use sits on one account with one provider. When that access got suspended in June nobody has a second way in, my team did the tickets by hand for two days”*
>
> **INDIVIDUAL CONTRIBUTOR**
> CXM, Corporate and back-office functions

> *“I score that section down for one reason only, nobody wrote down what old ops team use to patch by hand before the monitoring agent came”*
>
> **INDIVIDUAL CONTRIBUTOR**
> ITO, Technology engineering and operations

5.2 Ratings Related to Third-Party Work Done by Hand

Ratings indicated gaps in retaining manual capability, knowing where informal corrections had occurred, and replacing those corrections with explicit controls. Each of the three indicators had a mean below the midpoint of the 10-point scale. They assess different aspects of the delivery team's preparation for third-party problems and are considered separately after the comparison in Table 5.3.

Table 5.3. Ratings related to third-party work done by hand



Indicator	Mean	Distressed ratings	r, cancellation
C190 Retention of the Ability to Perform Manual Workarounds	4.66	249 of 808 (30.82%)	−0.65
C191 Knowledge of Informal Third-Party Corrections	4.61	262 of 792 (33.08%)	−0.66
C192 Replacement of Informal Corrections With Explicit Controls	4.58	267 of 796 (33.54%)	−0.64

Indicator	r, customer satisfaction	r, intent to leave
C190 Retention of the Ability to Perform Manual Workarounds	0.72	−0.67
C191 Knowledge of Informal Third-Party Corrections	0.61	−0.66
C192 Replacement of Informal Corrections With Explicit Controls	0.71	−0.65

Note: Means are respondent-level arithmetic means, with bases of 808, 792, and 796 for C190, C191, and C192, respectively. Distressed ratings are ratings of 3 or below. The coefficient r measures the direction and strength of a linear association, from −1 to +1. Cancellation correlations use 150 engagements and customer-satisfaction correlations use the 94 active engagements. Intent-to-leave correlations use paired respondent answers: 757 for C190, 746 for C191, and 748 for C192. Related ratings do not establish the occurrence of third-party work done by hand. Each indicator's separate discussion below also reports its relationship with engagement indicator mean.

A 95% confidence interval (CI) expresses the uncertainty around an estimated association under the stated method. Cluster-bootstrap intervals resample whole engagements to retain the grouping of respondents within engagements. The intervals below are existing reported results. Correlations with engagement indicator mean exclude the indicator being tested to avoid including it in both quantities.

5.3 Retention of the Ability to Perform Manual Workarounds

Retained manual capability received 249 distressed ratings among 808 answers (30.82%), indicating gaps in the delivery team's reported ability to perform work by hand when third-party systems failed. The mean rating was 4.66 and the median was 5.0. Respondents on canceled engagements reported a mean of 3.2 (n = 311). These ratings concern retained capability, not the occurrence count for third-party work done by hand.

Higher capability ratings were associated with a lower cancellation rate (r = −0.65, 95% CI −0.73 to −0.54, n = 150 engagements), higher active-engagement customer satisfaction (r = 0.72, 95% CI 0.60 to 0.80, n = 94 engagements), and a higher engagement indicator mean calculated without this indicator (r = 0.88, 95% CI 0.84 to 0.91, n = 150 engagements). Higher ratings were also associated with lower intent to leave (r = −0.67, cluster-bootstrap 95% CI −0.73 to −0.62, n = 757 respondents). The engagement mean was distressed on 38 of the 150 engagements.

The indicator-specific compilation contained 43 written answers about manual capability, with 22 negative and 21 positive sentiment classifications. Respondents described both retained skills used during outages and capability lost when experienced delivery team members left without documenting or passing on the work. The existing written-sentiment analysis detected no statistically significant relationship with the core measures. Retaining a capability and relying on undocumented individual knowledge remain distinct aspects of the reported experience.

5.4 Knowledge of Informal Third-Party Corrections

Knowledge of where informal third-party corrections had occurred received 262 distressed ratings among 792 answers (33.08%). The mean rating was 4.61 and the median was 5.0. Respondents on canceled engagements reported a mean of 3.1 (n = 305). This indicator concerns knowledge of earlier manual practices rather than the delivery team's present ability to perform them.

Higher knowledge ratings were associated with a lower cancellation rate (r = −0.66, 95% CI −0.74 to −0.56, n = 150 engagements), higher active-engagement customer satisfaction (r = 0.61, 95% CI 0.47 to 0.73, n = 94 engagements), and a higher engagement indicator mean calculated without this indicator (r = 0.87, 95% CI 0.82 to 0.90, n = 150 engagements). They were also associated with lower intent to leave (r = −0.66, cluster-bootstrap 95% CI −0.71 to −0.60, n = 746 respondents). The engagement mean was distressed on 39 of the 150 engagements.

Of the 21 indicator-specific written answers, 20 had negative sentiment. The accounts described undocumented corrections, departures of people who knew them, and repeated rediscovery of the same correction when a third-party problem recurred. The existing written-sentiment analysis detected no statistically significant relationship with the core measures.

The accounts illustrate the knowledge gap without measuring how often it occurred across all engagements.

5.5 Replacement of Informal Corrections With Explicit Controls

Replacement of informal third-party corrections with explicit checks received 267 distressed ratings among 796 answers (33.54%). The mean rating was 4.58 and the median was 5.0. Respondents on canceled engagements reported a mean of 3.1 (n = 302). The indicator distinguishes knowing about a correction from incorporating it into a control when agents assumed the work.

Higher replacement ratings were associated with a lower cancellation rate (r = −0.64, 95% CI −0.73 to −0.54, n = 150 engagements), higher active-engagement customer satisfaction (r = 0.71, 95% CI 0.59 to 0.80, n = 94 engagements), and a higher engagement indicator mean calculated without this indicator (r = 0.88, 95% CI 0.84 to 0.91, n = 150 engagements). They were also associated with lower intent to leave (r = −0.65, cluster-bootstrap 95% CI −0.71 to −0.59, n = 748 respondents). The engagement mean was distressed on 40 of the 150 engagements.

Nine indicator-specific written answers addressed control replacement: seven had negative sentiment, one positive, and one neutral. They described corrections not incorporated into explicit checks when agents assumed tasks. The small base supports illustration rather than a frequency estimate of the incidents described. The structured ratings remain respondents' assessments of control replacement rather than an independent control test.

Manual Alteration by Cancellation Status and Engagement Indicator Mean

Manual alteration was reported across the full range of engagement indicator means, although four methods were more common on canceled engagements and engagements with lower indicator means. Respondents on all 56 canceled engagements described manual alteration of financial institution processes and AI outputs. They described manual alteration of financial-institution-provided data and reporting alteration on 55 of those engagements, and third-party work done by hand on 29.

The observed cancellation rate was higher among engagements on which respondents reported the four most common methods than among engagements with no such report. For AI-output manual alteration, cancellation was recorded on 45.90% of the 122 engagements with a report and on none of the 28 without one. The corresponding contrast for third-party work done by hand was smaller: 39.19% of 74 engagements with a report, compared with 35.53% of 76 without one. Absence of a report does not independently establish absence of the practice.

Table 5.4. Cancellation by the presence of each manual alteration method, 150 engagements

Method	Engagements with the method	Canceled	Engagements without the method	Canceled
Manual alteration of financial-institution-provided data	133	55 (41.35%)	17	1 (5.88%)
Manual alteration of financial institution processes	138	56 (40.58%)	12	0 (0.00%)
Manual alteration of AI outputs	122	56 (45.90%)	28	0 (0.00%)
Reporting alteration	126	55 (43.65%)	24	1 (4.17%)
Third-party work done by hand	74	29 (39.19%)	76	27 (35.53%)

Note: Cancellation was recorded on 56 of the 150 engagements. Method presence means at least one qualifying report on the engagement. The study observed each engagement once. These comparisons do not establish whether manual alteration preceded or followed cancellation. They are descriptive contrasts, not forecasts of cancellation.

Every indicator quintile included reports of manual alteration on all 30 engagements. An indicator quintile is one of five groups of 30 engagements ordered by engagement indicator mean. Four methods were less common in higher quintiles. AI-output manual alteration was reported on all 30 engagements in the lowest quintile and 13 in the highest. Third-party work

done by hand showed no comparable ordered pattern, with 16 engagements in both the lowest and highest quintiles.

Table 5.5. Engagements with each manual alteration method, by quintile of the engagement indicator mean

Method	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Manual alteration of financial-institution-provided data	30 of 30	29 of 30	29 of 30	26 of 30	19 of 30
Manual alteration of financial institution processes	30 of 30	30 of 30	30 of 30	29 of 30	19 of 30
Manual alteration of AI outputs	30 of 30	29 of 30	27 of 30	23 of 30	13 of 30
Reporting alteration	30 of 30	29 of 30	26 of 30	25 of 30	16 of 30
Third-party work done by hand	16 of 30	17 of 30	13 of 30	12 of 30	16 of 30
Any of the five methods	30 of 30	30 of 30	30 of 30	30 of 30	30 of 30

Note: Quintile 1 contains the 30 engagements with the lowest engagement indicator mean and quintile 5 the 30 with the highest. Cells count engagements with at least one qualifying report of the method. Methods overlap. Rows are not additive. Indicator quintiles describe relative positions within this study set rather than validated categories of engagement performance.

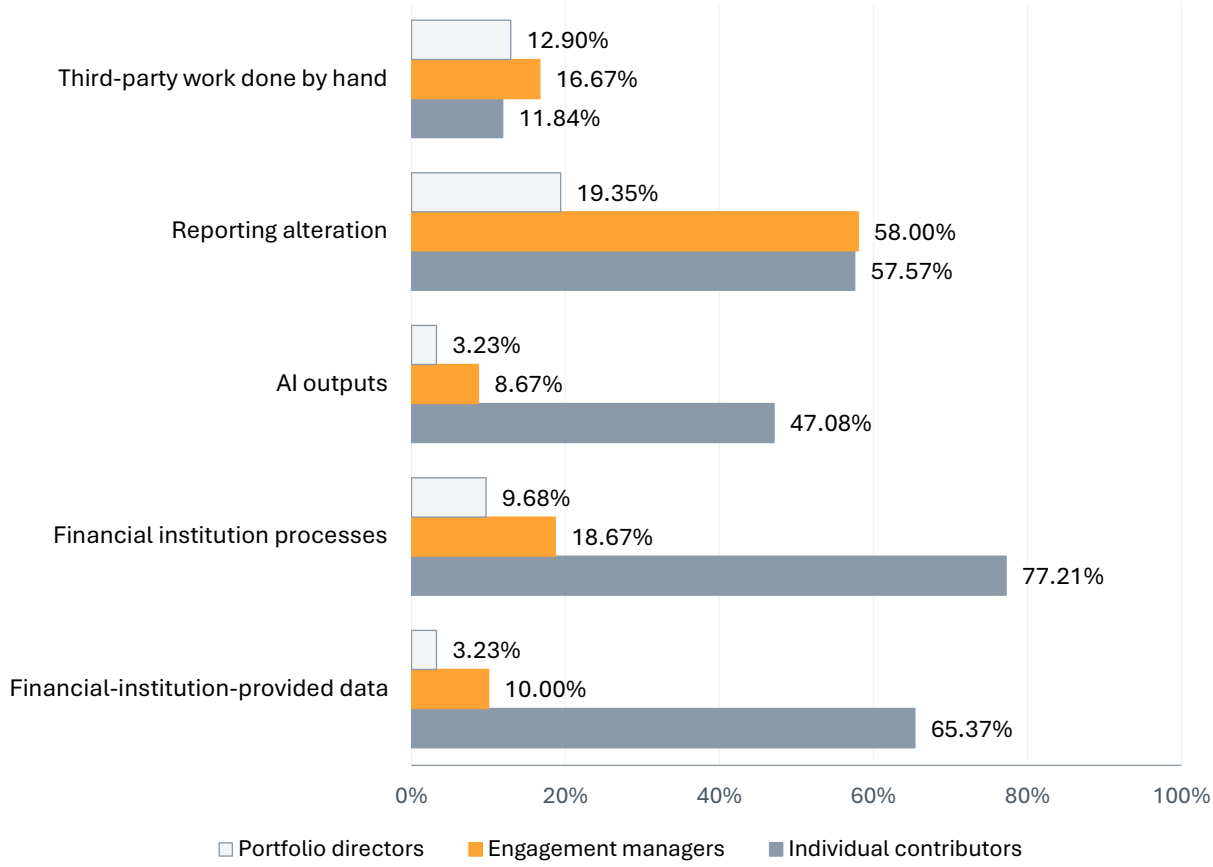
Distribution of Manual Alteration Across Roles

Individual contributors reported manual alteration more often than the other roles, with at least one method reported by 627 of 667 individual contributors (94.00%), 100 of 150 engagement managers (66.67%), and 11 of 31 portfolio directors (35.48%). Neither of the two customer management executives reported a method. These counts combine direct personal answers with written descriptions of practices on the engagement, so they do not measure personal performance of manual alteration uniformly across roles.

Individual contributors reported manual alteration of data, processes, and AI outputs more often than engagement managers. For process manual alteration, the counts were 515 of 667 individual contributors (77.21%) and 28 of 150 engagement managers (18.67%). Reporting alteration was described by similar shares of engagement managers and individual contributors: 87 of 150 (58.00%) and 384 of 667 (57.57%), respectively. Third-party work done by hand was described by 25 engagement managers (16.67%) and 79 individual contributors (11.84%). The comparisons describe reporting by role, not a direct allocation of responsibility for the work.

Figure 5.11. Respondents reporting each manual alteration method, by role

Manual alteration methods had different reported distributions across roles



Note: Shares use all respondents in each role: 667 individual contributors, 150 engagement managers, and 31 portfolio directors. Neither of the two customer management executives reported a method, and they are omitted from the figure. A report may concern personal action or a practice described on the engagement.

Table 5.6. Respondents reporting each manual alteration method, by role

Method	Individual contributors	Engagement managers	Portfolio directors
Manual alteration of financial-institution-provided data	436 of 667 (65.37%)	15 of 150 (10.00%)	1 of 31 (3.23%)
Manual alteration of financial institution processes	515 of 667 (77.21%)	28 of 150 (18.67%)	3 of 31 (9.68%)
Manual alteration of AI outputs	314 of 667 (47.08%)	13 of 150 (8.67%)	1 of 31 (3.23%)
Reporting alteration	384 of 667 (57.57%)	87 of 150 (58.00%)	6 of 31 (19.35%)
Third-party work done by hand	79 of 667 (11.84%)	25 of 150 (16.67%)	4 of 31 (12.90%)
Any of the five methods	627 of 667 (94.00%)	100 of 150 (66.67%)	11 of 31 (35.48%)

Note: A respondent is counted after a direct True answer or a written description of the method on the engagement. These sources do not uniformly establish personal performance of the method. Counts overlap across methods, and no report is not proof that a practice did not occur.

“

The vendor excluded misrouted regulated inquiries from the routing-accuracy dashboard by reclassifying them as client-side data errors after the fact, so the reported accuracy stayed above threshold while the actual misroutes kept mounting.”

ENGAGEMENT MANAGER

CXM, Client servicing and communications

1 Engagement Managers and Portfolio Directors

Reporting alteration accounted for most method reports from engagement managers. Of the 100 engagement managers reporting any method, 87 described reporting alteration, 15 reported data manual alteration, and 13 reported AI-output manual alteration. Among all 31 portfolio directors, six described reporting alteration and four described third-party work done by hand. The pattern concerns what these roles reported, rather than a verified division of work or knowledge.

> *“The vendor's portfolio reporting rolled this engagement's missed savings targets into a blended regional average, which kept the account's failure from showing clearly in any single dashboard view.”*
>
> **PORTFOLIO DIRECTOR**
>
> KPO, Corporate and back-office functions

“

Their bureau feed changed a field layout back in march and nobody flagged it... we found it when the challenger run wouldn't reconcile. Two days gone. And half of what we need from the Bureau portal has no real interface. You log in and export it by hand. That part is invisible here.”

INDIVIDUAL CONTRIBUTOR

KPO, Credit, underwriting and suitability

Reasons Given for Manual Alteration

Respondents most often linked manual alteration to unusable data, delayed financial institution decisions, unreliable AI outputs, and problems with third-party feeds and interfaces. The written answers describe reasons for changing or completing the work by hand. Table 5.7 maps those descriptions to the relevant survey questions and rated indicators. These mappings connect the accounts to Part 2 without establishing that the reported reasons caused the study outcomes.

Table 5.7. Survey subjects most frequently identified in written answers describing each manual alteration method

Method	Survey question or rated indicator	Answers	Engagements
Manual alteration of financial-institution-provided data	C094 Quality of Financial-Institution-Provided Data	341	103
Manual alteration of financial-institution-provided data	C108 Ability to Provision Access and Identities for Agents	92	61
Manual alteration of financial-institution-provided data	C096 Normalization of Financial Institution Data Across Source Systems	89	64
Manual alteration of financial institution processes	C109 Speed of Financial Institution Response to Issues Raised by the Engagement	229	108
Manual alteration of financial institution processes	C100 Completeness of Financial Institution Process Documentation	138	83

Table 5.7. Survey subjects most frequently identified in written answers describing each manual alteration method (continued)

Method	Survey question or rated indicator	Answers	Engagements
Manual alteration of financial institution processes	C105 Financial Institution Management of Its Side of the Engagement	94	62
Manual alteration of AI outputs	C157 Factual Reliability Before Human Review	321	106
Manual alteration of AI outputs	C159 Inconsistency of Output Across Repeated Runs	176	85
Manual alteration of AI outputs	C175 Stability Without Known Engagement-Side Changes	27	25
Reporting alteration	C157 Factual Reliability Before Human Review	53	37
Reporting alteration	C148 Human Review Capacity Against Agent Output Volume	8	8
Reporting alteration	C173 Explainability of Outputs to Audit Depth	7	7
Reporting alteration	C176 Diagnostic Value of Logs and Telemetry	7	7
Reporting alteration	C127 Quality Assurance for Nondeterministic Agent Output	7	7
Reporting alteration	C036 What stage is the engagement in?	7	7

Table 5.7. Survey subjects most frequently identified in written answers describing each manual alteration method (continued)

Method	Survey question or rated indicator	Answers	Engagements
Third-party work done by hand	C188 Usability of Third-Party Feed Data	36	27
Third-party work done by hand	C181 Advance Notice From Third-Party Providers	29	26
Third-party work done by hand	C187 Stability of Third-Party Interfaces and Data Structures	29	25

Note: Rows identify the three most frequent survey subjects in the coded written answers for each method, including every tie for third place. The method's own questions are excluded. Codes link descriptions to survey subjects and do not mean respondents used the formal indicator names. C036 is a stage question, not one of the 100 rated engagement indicators. An answer may concern several subjects. Counts overlap.

“

i cleaned the policy data again before load, we clean it every cycle and it comes back dirty again, i still don't have proper access to the source system so i pull what i can and patch the rest by hand.”

INDIVIDUAL CONTRIBUTOR

ITO, Client servicing and communications

1 Reasons for Manual Alteration of Financial-Institution-Provided Data

Data quality was the most frequently reported reason for manual alteration of financial-institution-provided data. It was identified in 341 of the 588 method-based written answers, covering 103 engagements. Borrowed and shared credentials were identified in 92 answers and normalization across source systems in 89. The accounts connected recurring correction with data quality and restricted access to the systems in which the problems originated.

Recurring correction in these accounts concerned the suitability of inputs for agent use and access to the systems in which errors originated. Correcting a supplied file does not by itself resolve the source-data or access problem. The reported reasons connect manual alteration to the data quality, accessibility, and normalization indicators examined in Part 2, “Financial-Institution-Provided Data.”

> *“I re-key the SSI details every time the layout shifts between deliveries, we still don't have working access to the source, so it's manual. Every time.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

“

It changes its mind between one pass and the next on the same input, this has occured on nearly every run this week, pressure from our own side is only to keep it moving, so I just pick the best version and send”

INDIVIDUAL CONTRIBUTOR

ITO, Technology engineering and operations

2 Reasons for Manual Alteration of Financial Institution Processes

Delayed decisions were the most frequently reported reason for process manual alteration. The financial institution's speed of response was identified in 229 of the 857 method-based written answers, covering 108 engagements. Process documentation was identified in 138 answers and financial institution management and decision authority in 94. These accounts describe delivery teams changing the work when the decisions or documented procedures needed to perform it were unavailable.

A changed process may allow work to continue while leaving its authoritative definition or the required financial institution decision unresolved. These accounts connect manual alteration to process ownership and authorization, not simply to the frequency of intervention. Part 2 distinguishes process documentation from the financial institution's decision authority and response speed. The accounts here do not establish that every process change lacked approval or was improper.

> *"Process never written down anywhere, everyone do it different way. I ask questions and nothing come back so I change step myself to keep work moving."*
>
> **INDIVIDUAL CONTRIBUTOR**
> CXM, Corporate and back-office functions

> *"What they gave us was a slide, not a procedure. And change control cant keep up with the speed we work at so I just adjust it"*
>
> **INDIVIDUAL CONTRIBUTOR**
> CXM, Corporate and back-office functions

> *"Last month the Client raise a request that need urgent proccess change but formal Change board only meet monthly, so I make the update myself and inform team lead after"*
>
> **INDIVIDUAL CONTRIBUTOR**
> KPO, Credit, underwriting and suitability

> *The change process was built for quarterly release, not for agent making decision every day… so when servicing rule need small update I please make the same myself rather than wait weeks for formal sign off"*

INDIVIDUAL CONTRIBUTOR

BPO, Payments, cards and disputes

3 Reasons for Manual Alteration of AI Outputs

Unreliable output before human review was the most frequently reported reason for AI-output manual alteration. It was identified in 321 of the 429 method-based written answers, covering 106 engagements. Inconsistency across repeated runs was identified in 176 answers. The separate reasons compilation recorded instructions to continue processing outputs without raising the problem in 16 answers from 12 respondents on 10 engagements. This differs from the method-based count of 14 answers on nine engagements reported earlier. The counts remain separate because the coverage of that theme has not been reconciled.

The reported correction and selection of outputs place human review within the work required to produce an acceptable result. A corrected output cannot by itself demonstrate the factual reliability or repeatability of the agent's uncorrected work. Whether the human contribution was recorded and visible in performance reporting is examined separately below. The two escalation counts remain descriptions from their respective compilations.

> *"yesterday i ran same question twice and bot gave different policy answer both time, nothing else changed. and i had to correct it before new hire saw wrong one."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Corporate and back-office functions

> *"Nothing about it is repeatable, i get different Output on same Input twice in one day and Change process assumes normal release cycle not agent running every day"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Trade, settlement and reconciliation

> *"The tool gives a different attribution breakdown each time I rerun it... consistency is the one thing it cannot do and my own leadership does not want that raised with the client"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Trade, settlement and reconciliation

“

The vendor's delivery dashboard reported reporting-rule coverage as a percentage of rules migrated into the agent, but counted a rule as migrated the moment a draft version was uploaded, regardless of whether an SME had validated it, so coverage looked near complete while unvalidated rules kept producing quiet errors downstream.”

INDIVIDUAL CONTRIBUTOR

KPO, Regulatory, audit, tax and legal

4 Reasons for Reporting Alteration

Factual reliability before human review was the most frequently identified additional subject in accounts of reporting alteration, appearing in 53 answers on 37 engagements. Most accounts described how reports were altered rather than supplying a separate reason for the practice. Table 5.2 distinguishes those reporting methods from the additional survey subjects summarized here.

The distinction between a reporting method and a reason for using it limits interpretation of these counts. Accounts of corrected work attributed to agents connect reporting alteration to the visibility of human effort, but identifying such a method does not independently establish the respondent's or vendor's motive. The relevant comparison is between the work described and what its reported performance represented.

> *“The vendor's onboarding metrics counted a task as agent-handled even when a human quietly redid it afterward because the agent's output was unusable, which overstated automation coverage.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Corporate and back-office functions

“

The feeds themselves have settled — the supplier registry and sanctions data have not moved under us once this year, and our contracts with those providers now say something sensible about notice periods, which was not true in January. Where we have not moved is our own map of who sits behind them. We know our direct providers; we do not honestly know their sub processors, and the truthful answer on the fourth party question is that we never finished the exercise. Nor could we run the onboarding checks by hand for long if one of them went dark — the people who used to do that work are doing something else now.”

ENGAGEMENT MANAGER

BPO, Corporate and back-office functions

5 Reasons for Third-Party Work Done by Hand

Unusable feeds, unstable interfaces, and insufficient notice of provider changes were the reported reasons for third-party work done by hand. Feed-data usability was identified in 36 of the 113 method-based written answers, covering 27 engagements. Interface and data-structure stability and advance notice from third-party providers were each identified in 29 answers. The accounts describe manual work needed when the delivery team could not rely on the third-party connection or its data.

These accounts place part of the effort needed to use third-party services within the delivery team's own work. They identify a potential continuity dependency on the people who know how to complete that work by hand. The related ratings of retained manual capability and replacement of informal corrections remain separate from counts of reported occurrence.

“

i type the return into Submission portal by hand every time, agent cannot use a screen made for a person. give us a proper interface, and please tell us before they change the format, that is what has to happen”

INDIVIDUAL CONTRIBUTOR

KPO, Regulatory, audit, tax and legal

Disclosure of Manual Alteration

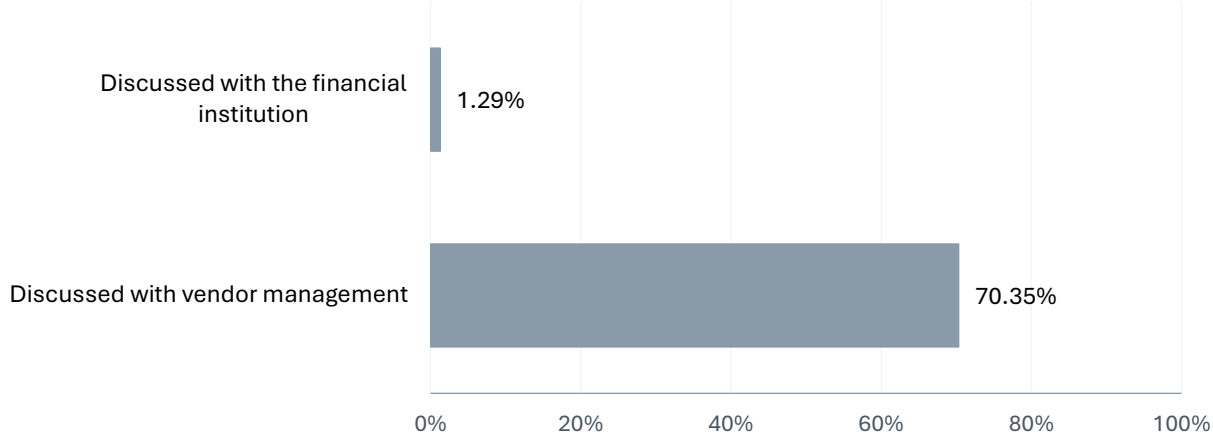
Answers about discussion of manual alteration with vendor management were much more frequently affirmative than answers about discussion with the financial institution. The overall contrast pools answers to separate questions and does not describe a matched comparison of the same people.

> *"My own management told me to keep it at working level and not put it in the client summary. So there are two versions of the issue list and neither one is complete."*
>
> **ENGAGEMENT MANAGER**
> KPO, Corporate and back-office functions

Across the three questions following personal manual alteration of data, processes, and AI outputs, 783 of 1,113 valid answers reported discussion with vendor management (70.35%). Discussion with the financial institution was reported in 13 of 1,004 valid answers (1.29%). These are pooled answers across three questions, not unique respondents or separate incidents. The two totals have different answering bases and do not form a matched comparison for every respondent.

Figure 5.12. Reported discussion of manual alteration with vendor management and the financial institution, pooled answers across three questions for each

Discussion with vendor management was reported more often than discussion with the financial institution



Note: Numerators count True answers and denominators count True or False answers: 783 of 1,113 for vendor management and 13 of 1,004 for the financial institution. A person may answer more than one question. The answers describe what respondents reported that they or their management discussed, rather than establishing the financial institution's knowledge.

Written descriptions also distinguished vendor management disclosure from financial institution disclosure, but the evidence was uneven across methods. Among the 2,410 method-based written answers, 309 from respondents on 119 engagements stated that vendor management had been told. Financial institution nondisclosure was described in 106 answers on 51 engagements, while disclosure was described in 26 on 19 engagements. Another 44 answers on 31 engagements stated that nobody had been told. These themes overlap and are not an exhaustive accounting of disclosure on all engagements.

Table 5.8. Disclosure of each manual alteration method

Method	Discussed with internal management	Discussed with the financial institution	Told vendor management
Manual alteration of financial-institution-provided data	292 of 401	3 of 366	105 (69)
Manual alteration of financial institution processes	341 of 459	9 of 414	162 (87)
Manual alteration of AI outputs	150 of 253	1 of 224	51 (38)
Reporting alteration	Not asked	Not asked	0 (0)
Third-party work done by hand	Not asked	Not asked	6 (6)

Table 5.8. Disclosure of each manual alteration method / Panel B

Method	Financial institution was told	Financial institution was not told	Told nobody
Manual alteration of financial-institution-provided data	8 (7)	41 (28)	19 (16)
Manual alteration of financial institution processes	17 (14)	52 (32)	13 (12)
Manual alteration of AI outputs	3 (2)	17 (14)	12 (11)
Reporting alteration	0 (0)	0 (0)	0 (0)
Third-party work done by hand	0 (0)	1 (1)	0 (0)

Note: The first two columns report True answers among valid True or False answers to the disclosure questions following personal manual alteration of data, processes, and AI outputs. They are question-specific answering bases. The final four columns count method-based written answers, with engagements in parentheses. Themes and methods overlap. Zero means no identified statement in that coded set, not verified absence of disclosure. Reporting-alteration answers concerned the vendor's reports and supplied little evidence about the respondent's own disclosure.

“

we have told the client directly that we alter their data before it reaches the model, and told our own management as well, also I am careful how that gets phrased since altering anything on their side sits close to a line we are not meant to cross without sign-off.”

ENGAGEMENT MANAGER

CXM, Financial crime and client onboarding

1 Disclosure to Vendor Management

Most respondents answering the internal-disclosure questions reported discussing manual alteration with vendor management, although discussion was less common for AI outputs than for data or processes. True answers numbered 292 of 401 for data (72.82%), 341 of 459 for processes (74.29%), and 150 of 253 for AI outputs (59.29%). Each percentage uses that question's own valid-answer base.

Internal disclosure of data or process manual alteration was associated with a lower cancellation rate, a higher engagement indicator mean, and lower intent to leave at engagement level. The data comparison for engagement indicator mean did not retain statistical significance under the Kruskal–Wallis check, a rank-based comparison of groups, and its customer-satisfaction comparison could not be estimated. AI-output disclosure had no statistically significant engagement-level association, although two respondent-level checks accounting for clustering within engagements did. The differences limit a single interpretation of disclosure as beneficial.

> *"I revert back the numbers to match the Statement and keep the same in my own tracker… I told my Lead, we never discuss about it with the client."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

1.1 Internal Disclosure of Data Manual Alteration

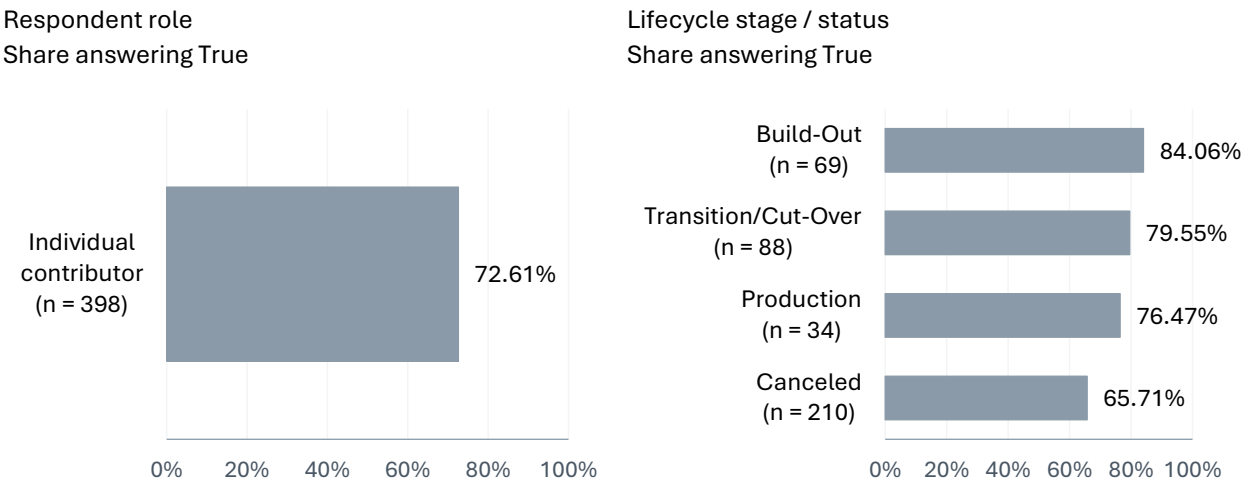
Indicator C200 · I discussed the data alteration with internal management · Full statistical analysis ›

Discussion of data manual alteration with vendor management was reported by 292 of 401 respondents answering True or False (72.82%), while 109 answered False (27.18%). The reported share showed no statistically significant differences across the groups of study classifications after Holm correction. That correction adjusts comparisons within a defined family of tests to limit false-positive conclusions from multiple testing.

Figure 5.13. Share answering True to "I discussed the data alteration with internal management", by respondent role and engagement status

Discussion of data manual alteration with vendor management was reported by 292 of 401 respondents answering True or False (72.82%), while 109 answered False (27.18%).

Engagement managers: 3 of 3 answered True. This small answering base does not support a stable management estimate.



Note: True or Yes answers: 292 of 401 answers (72.82%).

n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

Group shares and answering bases: Individual contributor (n = 398): 72.61%. Engagement manager (n = 3): 3 of 3 answered True. Build-Out (n = 69): 84.06%. Transition/Cut-Over (n = 88): 79.55%. Production (n = 34): 76.47%. Canceled (n = 210): 65.71%.

The cancellation rate was lower among engagements assigned True on internal data disclosure: 39.77% of 88 engagements, compared with 64.00% of 25 assigned False (chi-square 4.61, df = 1, p = 0.032, n = 113 engagements). Engagement indicator mean (Welch F 6.49, p = 0.014) and intent to leave (Welch F 9.59, p = 0.003) also differed, each on 113 engagements. The indicator-mean result did not retain statistical significance under the

Kruskal–Wallis check. Active-engagement customer satisfaction could not be compared because fewer than two answer categories included at least 10 engagements.

The 210 indicator-specific written answers included 172 with positive sentiment and 37 with negative sentiment. Positive answers described telling a team lead or manager, sometimes while also stating that the financial institution was not told. Negative answers described telling nobody or nobody more senior than the respondent. The sentiment classification therefore concerns the subject of internal discussion and does not establish complete disclosure or adequate oversight.

> *I dont say to client, only to my Lead... takes time I dont have, and if I pick wrong duplicate nobody would know."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

> *And this happen almost every week, i catch it before it move on. i tell my manager same day, because of that it stop right there, no need to trouble client with it."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

> *This is routine for my team now, we get bad informations every week and nobody fixes the Source, I clean it myself and tell my Lead, not the client"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Corporate and back-office functions

1.2 Internal Disclosure of Process Manual Alteration

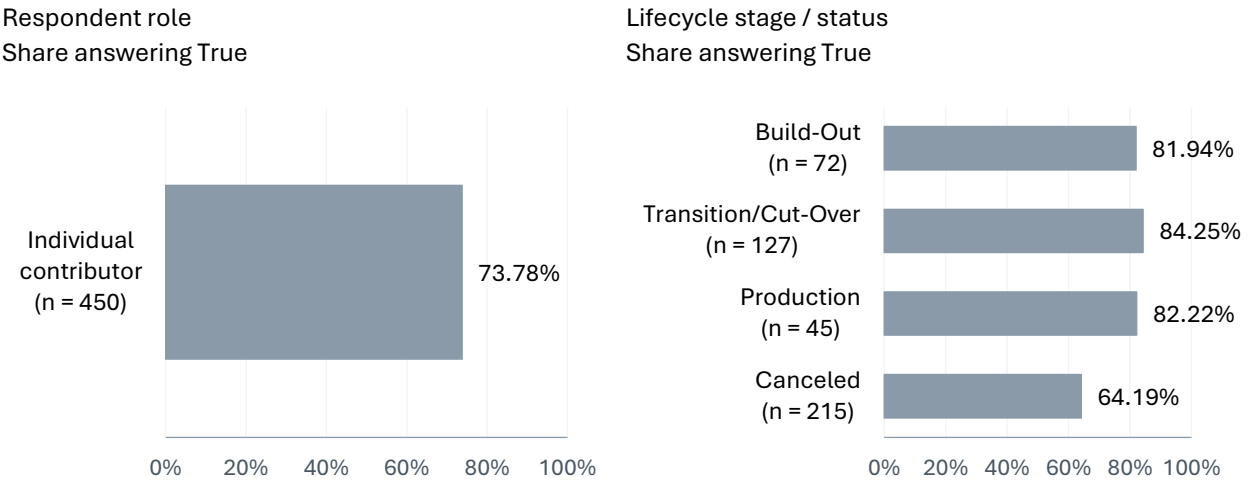
Indicator C206 · I discussed the process alteration with internal management · Full statistical analysis ›

Discussion of process manual alteration with vendor management was reported by 341 of 459 respondents answering True or False (74.29%), while 118 answered False (25.71%). No study classification family showed a statistically significant difference in the reported share after Holm correction.

Figure 5.14. Share answering True to "I discussed the process alteration with internal management", by respondent role and engagement status

Discussion of process manual alteration with vendor management was reported by 341 of 459 respondents answering True or False (74.29%), while 118 answered False (25.71%).

Engagement managers: 9 of 9 answered True. This small answering base does not support a stable management estimate.



Note: True or Yes answers: 341 of 459 answers (74.29%).

n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

Group shares and answering bases: Individual contributor (n = 450): 73.78%. Engagement manager (n = 9): 9 of 9 answered True. Build-Out (n = 72): 81.94%. Transition/Cut-Over (n = 127): 84.25%. Production (n = 45): 82.22%. Canceled (n = 215): 64.19%.

The cancellation rate was 35.00% among the 100 engagements assigned True on internal process disclosure and 65.22% among the 23 assigned False (chi-square 7.08, df = 1, p = 0.008, n = 123 engagements). The True group also had a higher engagement indicator mean (Welch F 12.68, p = 0.001) and lower intent to leave (Welch F 14.96, p < 0.001), each on 123 engagements. Active-engagement customer satisfaction could not be compared because fewer than two answer categories included at least 10 engagements.

The 257 indicator-specific written answers included 225 with positive sentiment and 29 with negative sentiment. Respondents described reporting changes to the team lead after making them. Some also stated that the financial institution was not told, distinguishing internal discussion from disclosure to the financial institution.

> *"Same workaround every week, we do the fix and backfill the paperwork after. And training the next new hire on it is getting old."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Trade, settlement and reconciliation

1.3 Internal Disclosure of AI-Output Manual Alteration

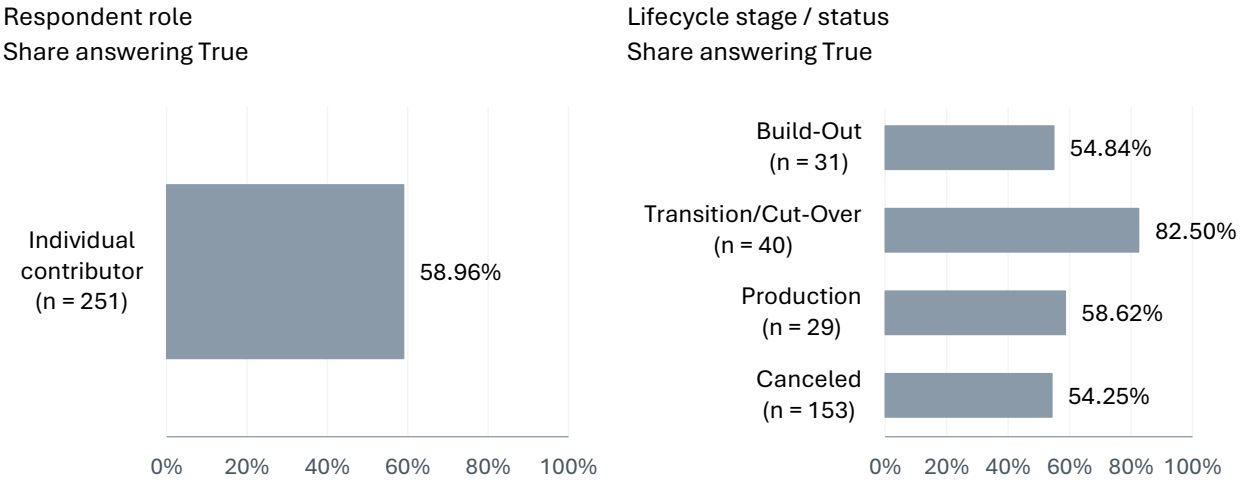
Indicator C212 · I discussed the AI output alteration with internal management · Full statistical analysis ›

Discussion of AI-output manual alteration with vendor management was reported less often than discussion of data or process manual alteration. True answers numbered 150 of 253 valid answers (59.29%). No statistically significant differences across study classification families were detected after Holm correction.

Figure 5.15. Share answering True to "I discussed the AI output alteration with internal management", by respondent role and engagement status

Discussion of AI-output manual alteration with vendor management was reported less often than discussion of data or process manual alteration.

Engagement managers: 2 of 2 answered True. This small answering base does not support a stable management estimate.



Note: True or Yes answers: 150 of 253 answers (59.29%).

n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

Group shares and answering bases: Individual contributor (n = 251): 58.96%. Engagement manager (n = 2): 2 of 2 answered True. Build-Out (n = 31): 54.84%. Transition/Cut-Over (n = 40): 82.50%. Production (n = 29): 58.62%. Canceled (n = 153): 54.25%.

Internal AI-output disclosure had no statistically significant engagement-level association with any of the four core measures. The existing tests reported chi-square 1.94 for cancellation (df = 1, p = 0.163, n = 93 engagements), Welch F 3.67 for active-engagement customer satisfaction (p = 0.063, n = 42 engagements), Welch F 3.13 for engagement indicator mean (p = 0.080, n = 93 engagements), and Welch F 2.32 for intent to leave (p = 0.131, n = 93 engagements). Two respondent-level checks accounting for clustering within engagements reached statistical significance for engagement indicator mean and intent to leave. That difference in analytical

unit does not establish an engagement-level association where the primary tests did not detect one.

The 95 indicator-specific written answers included 67 with positive sentiment and 26 with negative sentiment. Respondents who reported no disclosure described correcting outputs without raising the issue, sometimes because earlier escalation had not changed the situation.

> *This one i don't mention to anyone, not even my lead. Output changes between runs and i just pick version that looks right, nobody asks how it happened."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Corporate and back-office functions

> *i do not raise this to anyone... it happen most days, i kindly correct the output myself and send it, there is no other option for me"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Technology engineering and operations

> *Every person on client side do process different way, we chase same point on every status call and nothing move, raising ticket take longer than doing fix myself so most time I just fix it and tell my lead after"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Regulatory, audit, tax and legal

“

Last week the same paragraph came out different twice in a row on the same draft, consistency is the one thing it cannot do, and I catch it before it reaches an investor every time, that is simply the job and nobody logs it separately”

INDIVIDUAL CONTRIBUTOR

CXM, Client servicing and communications

2 Disclosure to the Financial Institution

Discussion with the financial institution was rarely reported in the three structured disclosure questions. True answers numbered three of 366 for data, nine of 414 for processes, and one of 224 for AI outputs. No engagement was assigned True under the plurality rule, so an engagement-level comparison could not be estimated. That assignment result does not mean no individual respondent reported disclosure.

> *I fix it quiet. And nobody official know except my team. Client never ask, I do not know if they hear. It happen every week, feels normal but not right."*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Corporate and back-office functions

> *I do this most days now, my manager know about it, please note we dont tell client, half the records missing values every batch and somebody has to catch that before it break something downstream,"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> ITO, Technology engineering and operations

2.1 Disclosure of Data Manual Alteration to the Financial Institution

Indicator C201 · I or my management discussed data alteration with the client · Full statistical analysis ›

Discussion of data manual alteration with the financial institution was reported in three of 366 valid True or False answers (0.82%). No engagement was assigned True under the plurality rule. The answers support a respondent-level description, while an engagement-level association with the core measures could not be estimated.

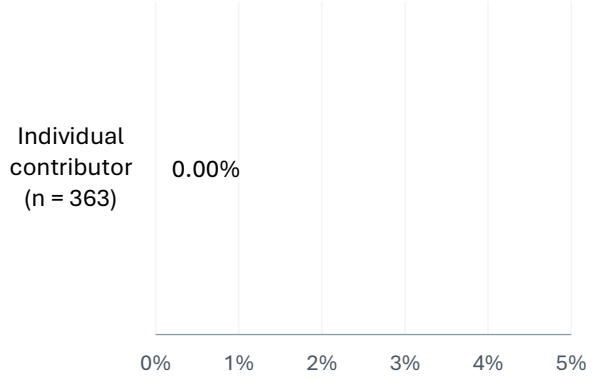
Figure 5.16. Share answering True to "I or my management discussed data alteration with the client", by respondent role and engagement status

Discussion of data manual alteration with the financial institution was reported in three of 366 valid True or False answers (0.82%).

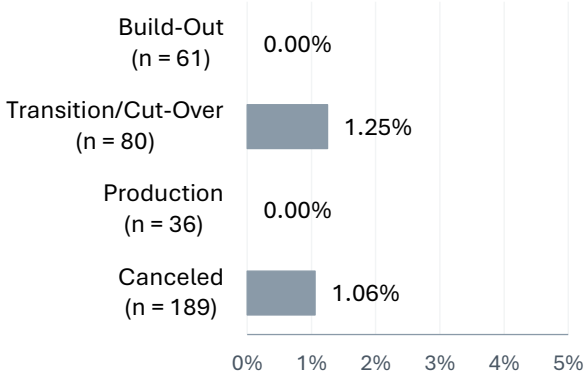
Detail scale: 0–5%. Exact answering bases and small-group qualifications are reported in the note.

Engagement managers: 3 of 3 answered True. This small answering base does not support a stable management estimate.

Respondent role
Share answering True



Lifecycle stage / status
Share answering True



Note: True or Yes answers: 3 of 366 answers (0.82%).

n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

Group shares and answering bases: Individual contributor (n = 363): 0.00%. Engagement manager (n = 3): 3 of 3 answered True. Build-Out (n = 61): 0.00%. Transition/Cut-Over (n = 80): 1.25%. Production (n = 36): 0.00%. Canceled (n = 189): 1.06%.

Of the 109 indicator-specific written answers, 96 had negative sentiment. At least 53 of those 96 stated that the financial institution was not told or that the manual alteration remained within the delivery team or vendor management. Eleven answers stated that the financial institution had been told, had requested correction, or had been shown the changes. These written-answer counts use a different source and definition from the direct True-or-False count.

> *what they did is send us mortgage files that dont reconcile, happens most days at this point, I tell my Lead, we dont tell client,"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Credit, underwriting and suitability

2.2 Disclosure of Process Manual Alteration to the Financial Institution

Indicator C207 · I or my management discussed process alteration with the client · Full statistical analysis ›

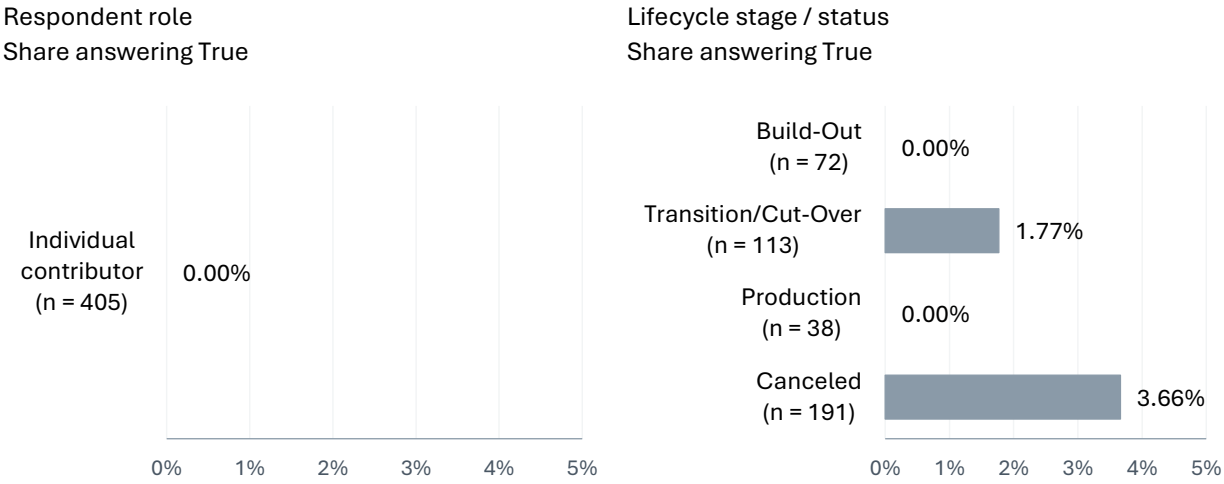
Discussion of process manual alteration with the financial institution was reported in nine of 414 valid True or False answers (2.17%). No engagement was assigned True under the plurality rule. With fewer than two answer categories containing at least 10 engagements, none of the four engagement-level comparisons could be estimated.

Figure 5.17. Share answering True to "I or my management discussed process alteration with the client", by respondent role and engagement status

Discussion of process manual alteration with the financial institution was reported in nine of 414 valid True or False answers (2.17%).

Detail scale: 0–5%. Exact answering bases and small-group qualifications are reported in the note.

Engagement managers: 9 of 9 answered True. This small answering base does not support a stable management estimate.



Note: True or Yes answers: 9 of 414 answers (2.17%).

n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

Group shares and answering bases: Individual contributor (n = 405): 0.00%. Engagement manager (n = 9): 9 of 9 answered True. Build-Out (n = 72): 0.00%. Transition/Cut-Over (n = 113): 1.77%. Production (n = 38): 0.00%. Canceled (n = 191): 3.66%.

Of the 124 indicator-specific written answers, 99 had negative sentiment. At least 57 of those 99 stated that the financial institution was not told or that the manual alteration remained within the delivery team or vendor management. Twelve answers stated that the financial institution had been told, and four described approval before a change was applied. These descriptions distinguish notification from prior approval, without independently verifying either.

> *“Same problem, they never gave us a real procedure, just a slide deck, so my team built our own steps, looped the client in as we went”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Corporate and back-office functions

> *“I tell Lead every time, we dont tell client, happens most days, informations we get is never enough to wait on approval, know which call to make most times now”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Corporate and back-office functions

2.3 Disclosure of AI-Output Manual Alteration to the Financial Institution

Indicator C213 · I or my management discussed AI output alteration with the client · Full statistical analysis ›

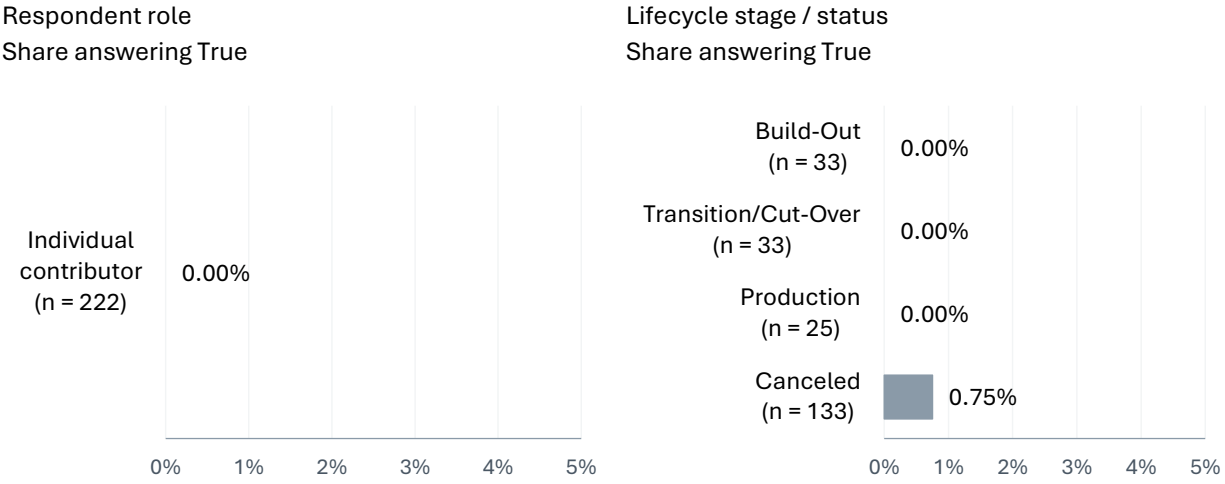
Discussion of AI-output manual alteration with the financial institution was reported in one of 224 valid True or False answers. No engagement was assigned True under the plurality rule, so an engagement-level association could not be estimated. The individual True answer remains part of the descriptive result.

Figure 5.18. Share answering True to "I or my management discussed AI output alteration with the client", by respondent role and engagement status

Discussion of AI-output manual alteration with the financial institution was reported in one of 224 valid True or False answers.

Detail scale: 0–5%. Exact answering bases and small-group qualifications are reported in the note.

Engagement managers: 1 of 2 answered True. This small answering base does not support a stable management estimate.



Note: True or Yes answers: 1 of 224 answers (0.45%).

n gives the numeric-answer base for each group. These are descriptive comparisons of reported answers.

Group shares and answering bases: Individual contributor (n = 222): 0.00%. Engagement manager (n = 2): 1 of 2 answered True. Build-Out (n = 33): 0.00%. Transition/Cut-Over (n = 33): 0.00%. Production (n = 25): 0.00%. Canceled (n = 133): 0.75%.

Of the 50 indicator-specific written answers, 43 had negative sentiment. At least 22 of those 43 stated that the financial institution was not told or that correction remained within the delivery team or vendor management. The single positive answer described telling both the financial institution and vendor leadership.

> *We just patch the breaks ourselves and keep the queue moving, nobody upstairs asks how they got cleared. And I couldn't tell you if the client ever sees the real touch count.”*

INDIVIDUAL CONTRIBUTOR

BPO, Trade, settlement and reconciliation

> *i tell team lead but not client… this happen maybe once a week, i fix informations in output before it go to review, otherwise sprint slip”*

INDIVIDUAL CONTRIBUTOR

ITO, Technology engineering and operations

> *When source-system feeds arrive late or malformed, the team logs a data exception, shares the informations with the client's platform owner, and agrees the correction before it is applied upstream. Nothing touches client data outside that channel.”*

PORTFOLIO DIRECTOR

ITO, Technology engineering and operations

3 Sources of the Disclosures That Occurred

Engagement managers supplied all 13 True answers on financial institution discussion across the three structured questions. They supplied 14 valid True or False answers in total, while individual contributors supplied 990 valid answers and none was True. These are answer counts across questions and do not establish the number of unique people who disclosed manual alteration.

Individual contributors supplied 1,175 of the 1,189 True answers on personal manual alteration of data, processes, and AI outputs. On discussion with vendor management, 769 of their 1,099 answers were True (69.97%). The contrast concerns personal-action and disclosure answers with different bases, rather than a matched sequence of incidents.

> *"We do flag it, both internally and to the client, but flagging does not stop us doing it, strictly speaking the account still needs the composites finished on time."*
>
> **ENGAGEMENT MANAGER**
>
> KPO, Trade, settlement and reconciliation

“

We just pick whichever record looks right and move on, raising this properly takes too long and nobody responds anyway. I don't discuss about what happens after with client, that's not for me, but this happens most weeks on this engagement.”

INDIVIDUAL CONTRIBUTOR

ITO, Technology engineering and operations

Recording of Manual Alteration

Respondents described manual alteration that was incompletely recorded and manual labor that did not appear in reported performance. In the method-based compilation, 107 written answers from 89 respondents on 56 engagements stated that manual alteration was not logged or only partly logged. Manual labor omitted from reported measures was described in 51 answers on 43 engagements, and undocumented practices dependent on an individual were described in 32 answers on 28 engagements. These accounts establish reported gaps in recording, not deliberate concealment in every case.

Table 5.9. Recording of each manual alteration method in the written answers

Method	Not logged or only partly logged	Manual effort omitted from reported measures	Undocumented practices dependent on individuals
Manual alteration of financial-institution-provided data	25 (22)	11 (11)	8 (8)
Manual alteration of financial institution processes	64 (40)	7 (7)	14 (12)
Manual alteration of AI outputs	19 (15)	10 (10)	4 (4)
Reporting alteration	2 (2)	27 (25)	0 (0)
Third-party work done by hand	2 (2)	4 (4)	9 (9)

Note: Cells count method-based written answers, with engagements in parentheses. An answer may describe several themes or methods. Rows and columns overlap. A zero identifies no coded statement in this set rather than verified absence of the practice. The table distinguishes recording gaps, omitted labor, and dependence on individual knowledge.

> *“the other thing is when the agent got a requirement wrong on a case, and it does, nobody above wanted to say it was ours. it went round for two weeks. my lead said just fix it and log it as data issue. so we did.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Credit, underwriting and suitability

> *Some of it makes it into the tracker, plenty does not, this is not every week, maybe once or twice a month, I share informations with my lead when it happen and we move forward”*

INDIVIDUAL CONTRIBUTOR

KPO, Credit, underwriting and suitability

1 Systems of Record

Records described by respondents were sometimes limited to team trackers, personal notes, or shared sheets rather than a complete engagement record. Process manual alteration accounted for the largest count of incomplete-recording descriptions: 64 written answers stated that a change was not logged or only partly logged. A personal or informal record could retain useful knowledge while still leaving the engagement's formal account incomplete.

> *"i tell my TL but not client, they dont ask anyway, there is no place to put rest of it, so it just disappear off record."*
>
> **INDIVIDUAL CONTRIBUTOR**
> CXM, Corporate and back-office functions

> *"i tell my TL, kindly he say just manage it... we do the same thing every week, i keep no record of the informations i change."*
>
> **INDIVIDUAL CONTRIBUTOR**
> CXM, Corporate and back-office functions

“

What we settle on in the end is a parallel note. Every synthesis the agents produce, one of my seniors reads it against the source before it goes over, and we hold a sheet of our own recording what changed and why. It is not part of any joint process, the client never sees the sheet. But when they come back three weeks later and ask why a figure moved, we can show them, and that is the only reason those conversations stay civil. It is a lot of hours for something that should be in the QA design itself.”

ENGAGEMENT MANAGER

KPO, Corporate and back-office functions

2 Visibility in Reported Performance

Respondents described performance reports attributing human-corrected work to agents while omitting the labor needed to produce the result. Human-corrected work counted as the agent's was described in 52 answers on 45 engagements. Manual labor absent from all reported measures was described in 51 answers on 43 engagements. These are overlapping descriptions, not a quantified estimate of overstatement. The study did not measure how far reported performance was overstated.

> *no one logs it as a defect, it's just cleanup we do every time we recieve a bad batch, so it never counts as extra work"*
>
> **INDIVIDUAL CONTRIBUTOR**
> BPO, Corporate and back-office functions

> *I don't report it as a fix, I just note it in our internal Tracker for the team... my manager sees it but nobody escalates, it just eats time every single week"*
>
> **INDIVIDUAL CONTRIBUTOR**
> BPO, Corporate and back-office functions

> *security approval takes months so we keep reconciling duplicates by hand meanwhile what goes out the other end always looks clean"*
>
> **INDIVIDUAL CONTRIBUTOR**
> ITO, Technology engineering and operations

> *The vendor's monthly NAV accuracy report averaged in reruns after a manual correction as if the agent had gotten it right the first time, hiding how often fund accountants had to step in."*
>
> **INDIVIDUAL CONTRIBUTOR**
> BPO, Trade, settlement and reconciliation

“

I feel bad marking everything low, it is mostly the training. They put me on contract review with two day orientation, nothing about payments regulation, nothing about what a clause actually mean, I learn from the senior girl sitting next to me. People here can do the work if somebody teach them”

INDIVIDUAL CONTRIBUTOR

KPO, Regulatory, audit, tax and legal

Regulatory and Legal Context Reported by Respondents

Knowledge of the regulatory context was uneven among the respondents who answered the question. Of 795 answers, 548 named at least one law, regulation, guidance document, or regulatory concept, collectively covering all 150 engagements. The other 247 answers (31.07%) were “I do not know.” All 247 came from individual contributors on 126 engagements. Coverage of every engagement by at least one named reference does not establish that every delivery team member knew the relevant obligations.

The references concerned model risk, anti-money laundering, outsourcing and third-party risk, data protection, operational resilience, and conduct. They combined laws, supervisory guidance, and broader concepts rather than a uniform set of verified legal obligations.

The study did not test legal applicability or compliance. These answers identify the regulatory references respondents believed were relevant to their engagements. They do not establish which requirements legally governed a particular practice or whether the financial institution or vendor complied with them.

> *“the issue is if anyone ever asks how the number was arrived at, the answer involves a decision that i made on a tuesday afternoon because nobody else would.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Corporate and back-office functions

“

Last month I needed access to the holdings extract for the pricing Module and it took nine days to come. That is the only reason data access got a 6 from me, it is not a comment on their people. Their subject matter experts answer me same day, and the process is run the same way in every region, so there are no surprises there. The 6 is about waiting for access, nothing more.”

INDIVIDUAL CONTRIBUTOR

ITO, Technology engineering and operations

1 Regulatory Items Named by Respondents

Model risk governance was the most frequently named regulatory reference, appearing in 219 answers on 58 engagements. Anti-money laundering and know-your-customer requirements ("AML and KYC") appeared in 192 answers on 48 engagements, and third-party and outsourcing risk guidance in 189 answers on 52 engagements. Respondents named 40 distinct references in total. Seven others appeared in 173 or 174 answers each, including Office of the Comptroller of the Currency (OCC) outsourcing guidance, supervisory letter SR 11-7, California and other state privacy laws, the General Data Protection Regulation (GDPR), the Digital Operational Resilience Act (DORA), European Banking Authority (EBA) outsourcing guidelines, and the European Union (EU) AI Act.

Table 5.10. Regulatory references most frequently named by respondents as relevant to the engagement

Reference named	Answers	Engagements
Model risk governance	219	58
"AML and KYC"	192	48
Third-party and outsourcing risk guidance	189	52
OCC outsourcing guidance	174	48
SR 11-7 model risk guidance	174	48
CCPA and state privacy laws	174	48
GDPR	173	47
DORA	173	47
EBA outsourcing guidelines	173	47
The emerging EU AI Act	173	47
Operational resilience requirements	140	40
MAS and local regulator outsourcing rules and local data protection laws	120	31

Note: The table selects the most frequently named references from the 40 identified in 548 answers that named at least one. An answer may name several references. Counts reflect respondent descriptions rather than verified legal applicability. CCPA means California Consumer Privacy Act. MAS means Monetary Authority of Singapore. Other abbreviations are expanded in the preceding paragraph. Labels identifying guidance, legislation, and broader concepts do not give them equivalent legal status.

All 150 engagement managers and all 29 portfolio directors who answered named at least one regulatory reference. Among the 615 individual contributors who answered, 247 reported "I do

not know” (40.16%). Individual contributors also reported manual alteration most often (Table 5.6). The juxtaposition identifies a difference in reported knowledge and practice by role, without establishing that individual contributors’ work breached an obligation.

2 Relationship to the Work Subject to Manual Alteration

The regulatory references named by respondents concerned work addressed by all five manual alteration methods. They included data and privacy, model outputs and validation, outsourced processes and dependencies, and decisions supported by those processes. This overlap establishes the relevance of regulatory knowledge to the reported work, rather than the legal applicability of every reference to every manual alteration.

Explicit descriptions of regulatory exposure were less frequent than descriptions of uncertainty about accountability. Sixteen written answers from 14 respondents on 12 engagements linked a practice, gap, or incident to a breach or risk of breach. Unclear liability or regulatory accountability for agent errors was described in 83 answers from 80 respondents on 57 engagements. These statements report respondents' assessments, not legal findings.

“

Credentials to the client Repository expire and nobody renews them, so I use a workaround login just to keep the files moving… it's the same thing every status call, we chase it and nothing changes”

INDIVIDUAL CONTRIBUTOR

BPO, Corporate and back-office functions

Regulatory and Legal Knowledge of the Delivery Team

Respondents' ratings indicated limited confidence in their own regulatory understanding, their management's understanding, and the financial institution's discussion of obligations, while training received still less favorable assessments. Each understanding or discussion indicator had a mean below 5.0. These are perceptions of preparation and dialogue rather than tests of legal knowledge or compliance.

Regulatory training received a mean rating of 3.6 (n = 798 respondents), with 385 distressed ratings (48.25%). The pattern connects the reported manual work to a preparation concern: respondents described working within a regulatory context that they did not consistently consider themselves well equipped to understand or apply.

> *"To be fair the hands on agent training was useful, I use it daily. The regulatory side they could not teach us, we learn that from the client's returns"*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> KPO, Regulatory, audit, tax and legal

“

We scrub the incoming Data clean every single cycle, as well as that it just turns up dirty again the very next time, so I just spend hours rebuilding same fields week after week”

INDIVIDUAL CONTRIBUTOR

ITO, Technology engineering and operations

1 Regulatory Understanding of the Respondent and Management

Respondents gave their own and their management's regulatory understanding mean ratings below 5.0, indicating limited confidence in both. Their own understanding was rated by 800 respondents and management's by 516. The different answering bases do not establish a matched comparison between the two ratings or explain why other respondents did not rate management.

Higher ratings of both forms of understanding were associated with a lower cancellation rate, higher customer satisfaction among active engagements, a higher engagement indicator mean, and lower intent to leave. Each indicator remains a separate reported perception. Similar associations from the same survey do not provide independent tests of legal capability.

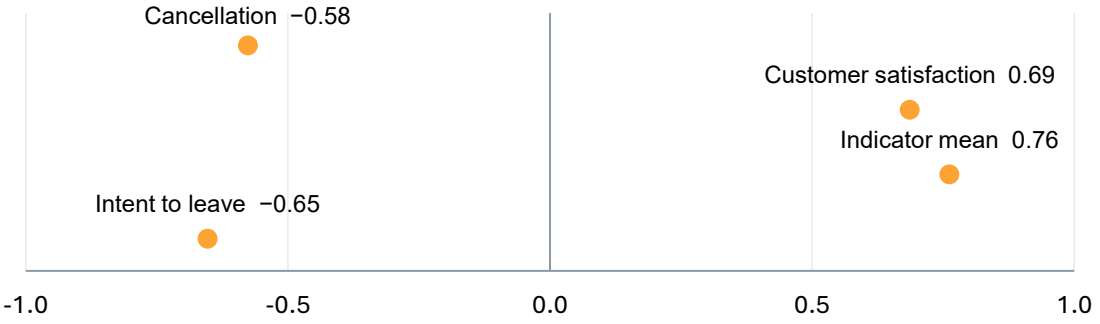
1.1 Respondent Understanding of the Governing Laws and Regulations

Indicator C057 · I fully understand how these laws, regulations, and similar affect the outsourcing engagement · Full statistical analysis ›

Respondents' self-assessed regulatory understanding received 228 distressed ratings among 800 answers (28.50%). The mean rating was 4.8 and the median was 5.0. Respondents on canceled engagements reported a mean of 3.5 (n = 304). Higher ratings were associated with a higher engagement indicator mean (r = 0.76, n = 150 engagements), lower intent to leave (r = −0.65, n = 754 respondents), a lower cancellation rate (r = −0.58, n = 150 engagements), and higher active-engagement customer satisfaction (r = 0.69, n = 94 engagements). The engagement mean for this indicator was distressed on 26 of the 150 engagements.

Figure 5.19. Respondent Understanding of the Governing Laws and Regulations: associations with the study measures

Respondent understanding of the governing laws and regulations: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.58 150 engagements p < 0.001
Customer satisfaction	r = 0.69 94 active engagements p < 0.001
Indicator mean	r = 0.76 150 engagements p < 0.001
Intent to leave	r = −0.65 754 respondents p < 0.001

Note: Survey question: "I fully understand how these laws, regulations, and similar affect the outsourcing engagement".

Distressed ratings: 228 of 800 answers (28.50%).

Group means and numeric-answer bases: Individual contributor (n = 623): 4.66. Engagement manager (n = 150): 5.21. Portfolio director (n = 27): 5.96. Build-Out (n = 164): 5.87. Transition/Cut-Over (n = 234): 5.39. Production (n = 98): 5.73. Canceled (n = 304): 3.48.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Eight indicator-specific written answers addressed regulatory understanding, supporting only a limited description. Seven identified regulatory knowledge learned on the engagement as something the respondent valued. The eighth described a delivery team encountering requirements as the agent flagged cases. Neither these accounts nor the self-ratings independently tested knowledge.

> *"Sharpened my sense of what regulatory accountability actually requires"*
>
> **INDIVIDUAL CONTRIBUTOR**
> ITO, Financial crime and client onboarding

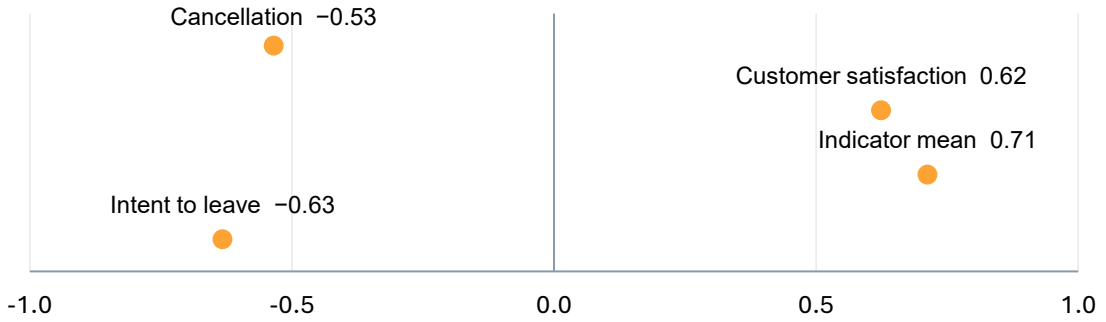
1.2 Management Understanding of the Governing Laws and Regulations

Indicator C058 · My management team fully understands these laws, regulations, and similar, and how they affect the engagement · Full statistical analysis ›

Management's regulatory understanding received 149 distressed ratings among 516 answers (28.88%). The mean rating was 4.9 and the median was 5.0. Respondents on canceled engagements reported a mean of 3.5 (n = 193). Higher ratings were associated with a lower cancellation rate (r = −0.53, n = 150 engagements), lower intent to leave (r = −0.63, n = 492 respondents), higher active-engagement customer satisfaction (r = 0.62, n = 94 engagements), and a higher engagement indicator mean (r = 0.71, n = 150 engagements). The engagement mean for this indicator was distressed on 31 of the 150 engagements. These findings describe the answering subset and do not establish what nonanswerers knew about management.

Figure 5.20. Management Understanding of the Governing Laws and Regulations: associations with the study measures

Management understanding of the governing laws and regulations: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.53 150 engagements p < 0.001
Customer satisfaction	r = 0.62 94 active engagements p < 0.001
Indicator mean	r = 0.71 150 engagements p < 0.001
Intent to leave	r = −0.63 492 respondents p < 0.001

Note: Survey question: "My management team fully understands these laws, regulations, and similar, and how they affect the engagement". Distressed ratings: 149 of 516 answers (28.88%).

Group means and numeric-answer bases: Individual contributor (n = 343): 4.72. Engagement manager (n = 150): 5.19. Portfolio director (n = 23): 6.35. Build-Out (n = 112): 6.10. Transition/Cut-Over (n = 155): 5.43. Production (n = 56): 6.04. Canceled (n = 193): 3.53.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

No written answer was coded to management's understanding of the governing obligations.

“

we do this every week now, the informations from client side never matches whats current, i just fix it and move on, nobody outside my team knows how much we touch before it reaches new hires.”

INDIVIDUAL CONTRIBUTOR

CXM, Corporate and back-office functions

2 Financial Institution Discussion of Obligations With the Delivery Team

Respondents' ratings indicated limited agreement that financial institution staff actively discussed regulatory and legal obligations with the delivery team. The mean was 4.8, and 231 of 801 ratings were distressed (28.84%). The answers describe perceived dialogue rather than the financial institution's own knowledge, documentation, or compliance.

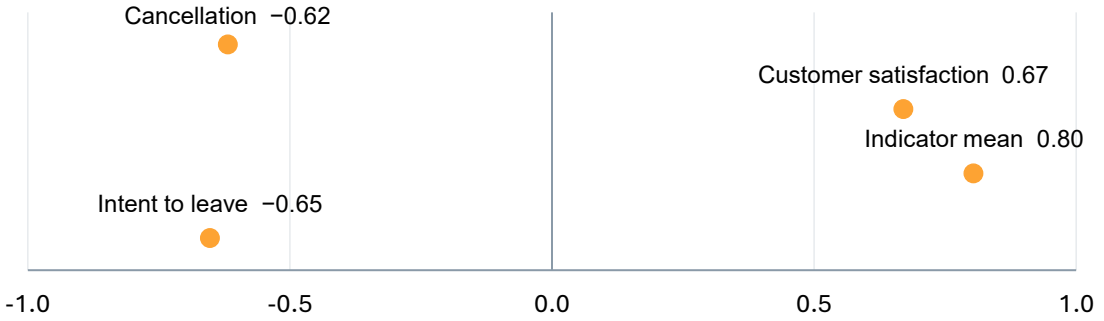
2.1 Discussion of Obligations With the Delivery Team

Indicator C059 · The customer's team actively discusses these laws, regulations, and similar with our team · Full statistical analysis ›

Discussion of obligations received a mean rating of 4.8 and a median of 5.0 (n = 801 respondents). Respondents on canceled engagements reported a mean of 3.4 (n = 305). Higher ratings were associated with a higher engagement indicator mean (r = 0.80, n = 150 engagements), a lower cancellation rate (r = −0.62, n = 150), higher active-engagement customer satisfaction (r = 0.67, n = 94 engagements), and lower intent to leave (r = −0.65, n = 752 respondents). Ratings were distressed in 231 answers (28.84%), and the engagement mean was distressed on 26 of the 150 engagements.

Figure 5.21. Discussion of Obligations With the Delivery Team: associations with the study measures

Discussion of obligations with the delivery team: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.62 150 engagements p < 0.001
Customer satisfaction	r = 0.67 94 active engagements p < 0.001
Indicator mean	r = 0.80 150 engagements p < 0.001
Intent to leave	r = −0.65 752 respondents p < 0.001

Note: Survey question: "The customer's team actively discusses these laws, regulations, and similar with our team".

Distressed ratings: 231 of 801 answers (28.84%).

Group means and numeric-answer bases: Individual contributor (n = 621): 4.63. Engagement manager (n = 150): 5.09. Portfolio director (n = 29): 6.34. Build-Out (n = 163): 5.66. Transition/Cut-Over (n = 232): 5.59. Production (n = 101): 5.68. Canceled (n = 305): 3.40.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

No written answer was coded to the financial institution's discussion of regulatory and legal obligations with the delivery team.

“

The help-desk dashboard reports agent deflection rate as tickets the agent handled without escalation, but tickets the agent closes and the same employee reopens within 24 hours are counted as a new, separate ticket rather than a reopened one, which hides the actual repeat-contact rate the deflection number is supposed to reflect.”

ENGAGEMENT MANAGER

CXM, Corporate and back-office functions

3 Regulatory Training

Regulatory training was rated less favorably than regulatory understanding or discussion of obligations. Its mean of 3.6 was the 19th-lowest among the 146 survey questions with a favorable or unfavorable direction. That comparison covers the broader set of directional survey questions, rather than only the 100 rated engagement indicators.

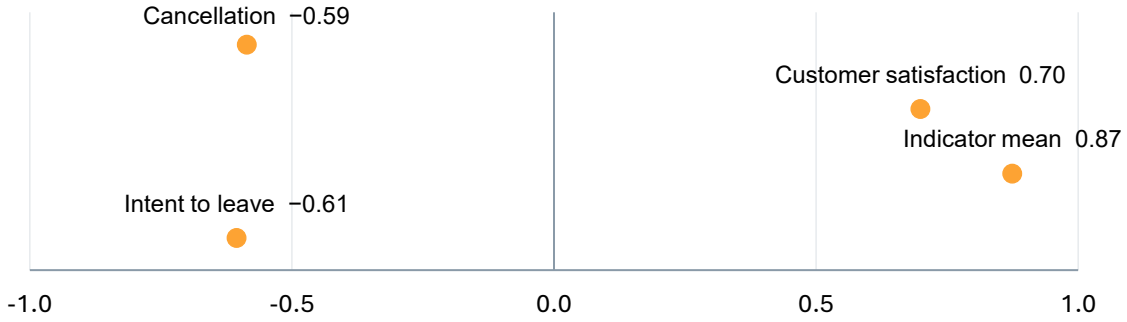
3.1 Adequacy of Regulatory Training

Indicator C060 · The training available to me effectively teaches me about these laws, regulations, and similar · Full statistical analysis ›

Nearly half of the regulatory-training ratings were distressed: 385 of 798 answers (48.25%). The mean rating was 3.6 and the median was 4.0. Respondents on canceled engagements reported a mean of 2.4 (n = 308). These assessments concern the perceived adequacy of available training, not an independent assessment of the instruction or respondents' legal knowledge.

Figure 5.22. Adequacy of Regulatory Training: associations with the study measures

Adequacy of regulatory training: associations with the study measures



Comparison	Estimate and analysis base
Cancellation	r = −0.59 150 engagements p < 0.001
Customer satisfaction	r = 0.70 94 active engagements p < 0.001
Indicator mean	r = 0.87 150 engagements p < 0.001
Intent to leave	r = −0.61 750 respondents p < 0.001

Note: Survey question: "The training available to me effectively teaches me about these laws, regulations, and similar".

Distressed ratings: 385 of 798 answers (48.25%).

Group means and numeric-answer bases: Individual contributor (n = 620): 3.48. Engagement manager (n = 150): 3.85. Portfolio director (n = 27): 4.22. Build-Out (n = 162): 4.37. Transition/Cut-Over (n = 234): 4.23. Production (n = 94): 4.48. Canceled (n = 308): 2.39.

r is Pearson's correlation coefficient, measuring signed linear association from −1 to +1. Whiskers show 95% confidence intervals where reported. Intent-to-leave intervals, where shown, use a cluster bootstrap. The bootstrap resamples engagements to preserve the grouping of respondent answers. Associations are observational.

Higher training ratings were associated with a higher engagement indicator mean (r = 0.87, n = 150 engagements) and higher customer satisfaction among active engagements (r = 0.70, n =

94 engagements). They were also associated with a lower cancellation rate (r = −0.59, n = 150 engagements) and lower intent to leave (r = −0.61, n = 750 respondents). The engagement mean for training was distressed on 64 of the 150 engagements.

All five indicator-specific written answers about regulatory training received the most negative sentiment score, −10. They described instruction that did not address the requirements relevant to the work. Four described learning instead from colleagues or financial institution material. The small number supports illustration of reported experience, rather than an estimate of how often these training gaps occurred.

> *“whoever trained us knows nothing about AML, consequently we learn the rules from the client's own manual.”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> BPO, Financial crime and client onboarding

> *“tool training was good, nobody taught us the suitability rules”*
>
> **INDIVIDUAL CONTRIBUTOR**
>
> CXM, Credit, underwriting and suitability

The reported experience indicates that manual alteration was widespread while disclosure, recording, and regulatory preparation were incomplete in important respects. Respondents described manual work that sustained delivery but was not consistently represented in engagement records or performance reports. Their accounts of limited financial institution disclosure and unfavorable training ratings make that limited visibility consequential for interpreting reported performance and the preparation of the delivery team. The study establishes neither the amount of performance overstatement nor legal noncompliance, and its cross-sectional associations do not establish that manual alteration caused cancellation.

“

I have always felt confident about how change is handled on this account. When the client's watchlist data arrives incomplete, the team raises it with the client's financial-crime unit, we discuss about the gap together, and any adjustment to triage process is agreed and logged before it is applied. We do not revert back to old workflows without that sign-off. Precedent across our sanctions engagements shows this is consistently how it works, five or six times a quarter at most.”

PORTFOLIO DIRECTOR

KPO, Financial crime and client onboarding

About the Author

Phil Hatch

Outsourcing and workforce development

Akholi author page

LinkedIn

ORCID 0000-0003-1709-8860

SSRN research paper

Phil Hatch is an outsourcing and workforce-development executive with more than 30 years of experience. He previously led the global outsourcing business of ManpowerGroup/Experis. His work includes service delivery, workforce strategy, and organizational development across industries and countries.

Hatch has also advised governments on labor and economic development, including workforce strategies intended to support investment and employment.

Following Russia's 2022 invasion of Ukraine, Hatch was asked by the U.S. Department of State and USAID to advise Ukraine on workforce and economic development. His work has addressed technology, outsourcing, skills development, and international investment in support of economic recovery.

Through Akholi, Hatch advises organizations and governments on outsourcing and the implications of AI for work and service delivery.

“

The vendor master data is a mess every time it lands with us. We clean it up each cycle and it just arrives dirty again. Their security team wont approve us a proper connection either, so I dont wait anymore, I just fix it myself.”

INDIVIDUAL CONTRIBUTOR

BPO, Corporate and back-office functions

About Akholi

Akholi, established in 2003, is a research and advisory firm focused on outsourcing, service delivery, and global capability centers. Its work includes the 2026 study of 150 agentic AI outsourcing engagements described in this report.

The study examines relationships among financial institution readiness, co-management, vendor capability, AI technology, the third-party ecosystem, and reported engagement outcomes. Its findings are observational and do not establish that any one category is the sole or primary cause of those outcomes.

Akholi applies this research through advisory and educational work with institutional leaders. Recommendations derived from the study require assessment against the circumstances and evidence of each engagement.

Akholi

PART 5 / 2026 RESEARCH STUDY

Manual Alteration by the Delivery Team

Part 5 examines five manual alteration methods, their distribution and stated reasons, disclosure, recording, and the regulatory and legal knowledge reported by delivery staff.

Phil Hatch | Akholi

Study ISBN 979-8-234-26882-2
SSRN DOI: https://doi.org/10.2139/ssrn.7164559
Figshare DOI: https://doi.org/10.6084/m9.figshare.34036890

Full study and supporting data

akholi.com/outsourcing-2026-agentic-ai-tier-1-banks

Version 1.8.2 | November 1, 2026

The evidence reflects respondents' reported experience. It is not an independent finding of regulatory breach.