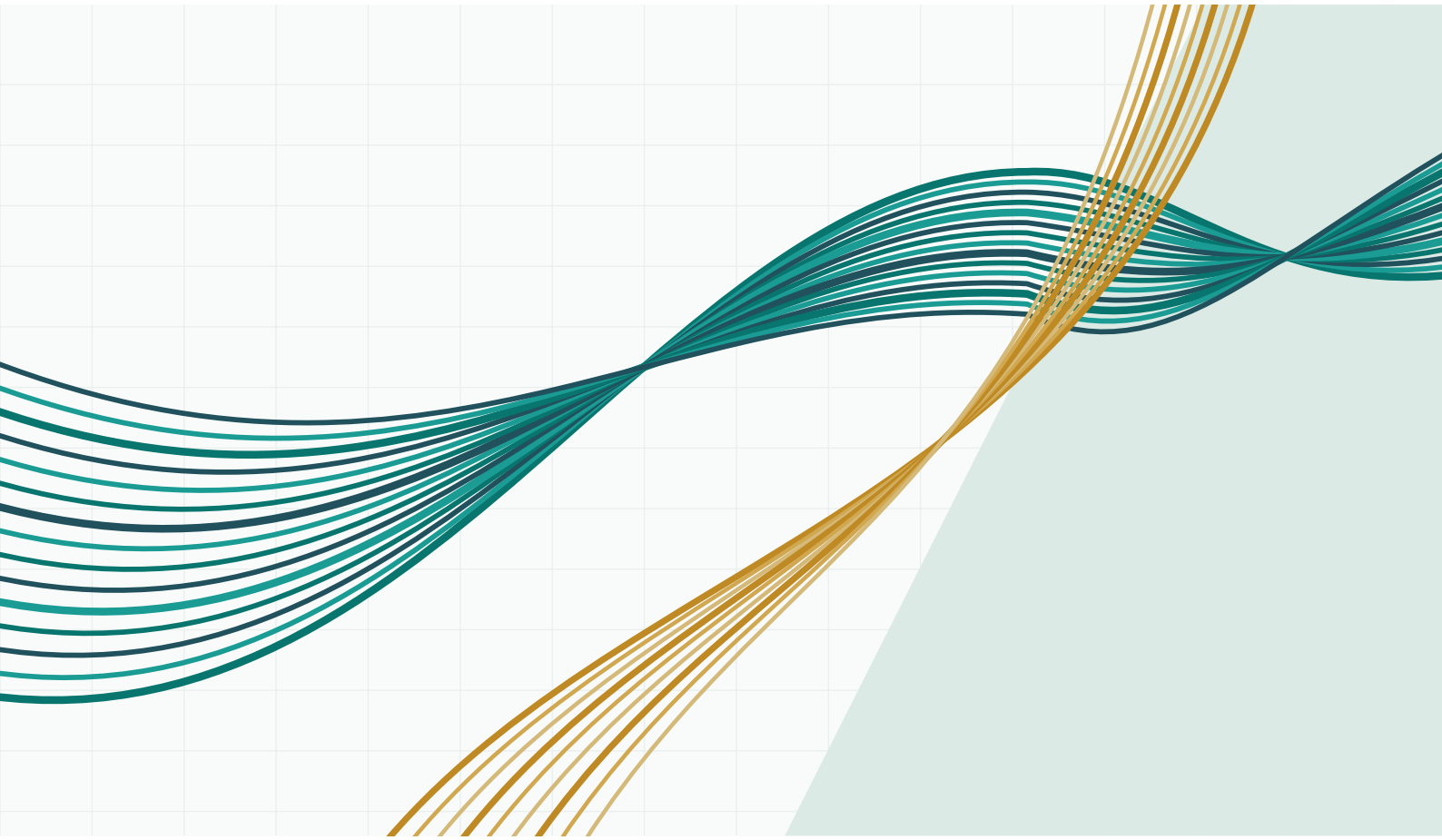
CAIA
CARIBBEAN ARTIFICIAL INTELLIGENCE ASSOCIATION
CAIA

RESEARCH & POLICY SERIES

REWEAVE

Reassessing jobs in the age of AI



CONSULTATION EDITION V0.1

November 2026

OVERVIEW

Reassessing jobs in the age of AI

Framework creator Adrian Dunkley, President, Caribbean AI Association

Publication Caribbean AI Association Research & Policy Series · Consultation edition v0.1 · November 2026

Abstract

When an organisation adopts AI, it can change which tasks people do, how they learn and which services receive their time. REWEAVE is a proposed method for reassessing one workflow at a time. Workers and managers identify what the service requires, remove unnecessary steps, allocate the remaining work, train for the redesigned role and earmark any capacity the change releases. A pilot compares staff effort, service quality, workload and demonstrated skill before management decides whether to expand. This paper draws on published research and gives worked examples, decision rules and a field protocol for Caribbean employers, worker representatives and training partners. The examples are hypothetical. A field comparison could show whether the framework improves outcomes relative to a tool-first rollout.

Executive decision

Authorise one bounded pilot with the people who do the work. Ask the process owner to show what the team stopped doing, which decisions remain with people, who can perform the redesigned job, and what the organisation did with the hours released. Require service, worker and safety checks before expanding.

Start with a frequent, reversible process. A claims intake queue, a hotel enquiries desk or a licensing application inbox offers clearer feedback than a whole-department redesign. Keep an approved fallback and record how to switch back.

Contents

Page	Section
3	What the evidence supports
4	The seven REWEAVE steps
5–6	Observe, simplify and allocate tasks
7	Train people and assign capacity
8–9	Design and assess a 90-day pilot
10–11	Three worked Caribbean applications
12	How to validate the framework
13–14	Task card and review worksheet
15–16	References

EVIDENCE

What the evidence supports

A job title can hide a mixture of routine transactions, judgement calls, mentoring and work that only appears when something goes wrong. Task-based economics distinguishes work displaced by automation from new work that technology creates [1]. Job-design research also treats autonomy, knowledge and relationships as features of work worth measuring [3]. REWEAVE applies these ideas to a decision a team can make and revisit.

Exposure shows where to look

The International Labour Organization (ILO) estimated that about one in four workers globally held an occupation with some exposure to generative AI in its 2025 analysis. The index measures potential at the level of tasks within occupations [4]. A separate 2024 ILO–World Bank study estimated that 26–38% of jobs in its Latin America and Caribbean analysis could be exposed. Its household-survey sample included the Dominican Republic, the only Caribbean island economy listed. The study identified gaps in digital access that may prevent workers from benefiting [5]. These studies use different methods and populations. Their figures cannot be read as an island estimate, a trend or an expected number of redundancies.

Results depend on the task and the worker

Study setting	Finding	Implication for a local pilot
Customer support at one company [6]	The study analysed 5,172 agents during a staggered rollout. It estimated that access raised issues resolved per hour by 15% on average, with larger gains for less experienced and lower-skilled agents.	Track outcomes by experience and case type where group sizes permit.
Professional writing experiment [15]	Participants using AI finished assigned writing tasks faster and received higher quality scores.	Test the actual documents and downstream review your service requires.
Consulting experiment [7]	Consultants benefited on some tasks. On a task outside the model's tested capabilities, participants using AI were less likely to reach the correct answer.	Use realistic exceptions when assessing a tool.
Experienced open-source developers [8, 9]	An early-2025 trial estimated slower completion with the tools tested. The authors described later estimates as weak evidence because of selection effects.	Count checking and repair. Retest when the tool, task or team changes.

OECD surveys in finance and manufacturing found that workers reported better outcomes where employers provided training and consulted workers. The association leaves room for other explanations [10]. The Caribbean AI Policy Roadmap makes education and upskilling a regional priority [11]. Workplace implementation still depends on the people, languages, equipment and operating conditions of each site.

FRAMEWORK

The seven REWEAVE steps

The method links two decisions. First decide whether to remove, simplify or keep each task. For tasks that remain, decide who or what will do them. Then test the resulting job, including its workload, authority and route for learning. A task can move to a machine while the job gets harder. Review both levels before signing off.

Step	Question	Evidence to carry forward
R · Review the work	What service must the workflow deliver, and what do people do to deliver it?	Observed task map, case mix, baseline and worker corrections.
E · Eliminate waste	Which steps can go, be combined or become simpler?	Dependency and control checks, plus a reversible removal trial.
W · Weigh who does what	Which person or system should do each remaining task?	Task decision, accountable owner, test cases and fallback.
E · Equip people	Can people perform the changed job and handle exceptions?	Paid practice, observed skills check and support plan.
A · Allocate capacity	If hours become available, who gets them and for what?	Time budget, named recipient, roster entry and outcome.
V · Verify results	Did the change improve the service without shifting unacceptable costs or risks?	Comparison, quality and workload checks, and an incident record.
E · Evolve roles	What should the organisation expand, revise or reverse?	Decision record, revised role description and next review date.

Appoint one operational owner with authority to pause. Include people who perform and receive the work, a worker representative where available, and whoever owns training and risk. In a small firm, one person can cover several roles. Give a second person authority to challenge the change and inspect the evidence.

> **Core requirement** A safety, data or accountability failure blocks the proposed delegation even if the time estimate is favourable. Workers need a channel to report failure, contest how their work is described and raise workload concerns.

REVIEW / ELIMINATE

Observe the work and remove steps

Record the job people perform

Choose an outcome with a clear start, finish and customer. For claims intake, start when a claim arrives and finish when an authorised officer confirms that it is ready for assessment. Track cases awaiting evidence separately from those an officer can complete. Record elapsed time for the customer and staff time for everyone who touches the case.

Observe a typical shift and a difficult one. Ask staff to show the handovers, phone calls, corrections and informal checks they rely on. Include people on different shifts and workers in junior or temporary roles. Let them amend the map before the process owner freezes the baseline. Sample edge cases such as a disconnected branch, a customer who prefers a local language, an incomplete form and an urgent request. Set sample size with the pilot lead after seeing how often each kind of case occurs.

Each task entry should name the trigger, output, recipient, frequency, staff minutes, downstream use, current control, decision authority and the training the task gives a newcomer. Record the source of each observation and what the team still needs to verify. A task can matter because it teaches judgement, even when its immediate output looks routine.

Try simplification before buying a tool

Ask who uses the output and what they do with it. Check audit, reporting and service dependencies with their owners. If two forms capture the same information, test whether an approved source record can feed both. If staff routinely send a manual reminder that an existing system already sends, check that the automated reminder reaches the right people and retains any required delivery record before withdrawing the duplicate.

Compare three options for any promising task: removal or simplification, a stable rule in existing software, and AI assistance. Use the cheapest option that meets the service and control requirements. When the team discovers an undocumented dependency, hold the removal, record the owner, and test a substitute with them.

Example A supervisor asks for a spreadsheet every Friday. Staff discover that finance uses only two columns to reconcile missing files. Agree a two-column extract with finance, keep the full record in the source system, and compare reconciliations for two cycles. The team can now decide whether the old sheet serves any remaining purpose.

WEIGH

Give each task an owner and a fallback

Decision	When it fits	What the owner records
Retire	Nobody needs the output after dependency and control checks.	Approver, removal date and restoration trigger.
Simplify	The outcome remains useful and the process has avoidable repetition or handovers.	New rule, affected teams and proof the control still works.
Keep human-led	The work needs discretion, relationships, judgement or consequential approval.	Named decision-maker, escalation threshold and trainee access.
Assist with AI	A person can check the output against a reliable source before using it.	Approved data, review steps, exceptions, maximum review time and fallback.
Automate within limits	The action is low consequence, tightly specified and reversible, and the team can detect errors. Use a deterministic rule where it suffices.	Boundaries, log, override, alert owner and rollback test.

For each task, first record what changes in the process: retire, simplify or keep as is. For every task that remains, record how the work will be done: human-led, AI-assisted or automated within limits. For example, a system can acknowledge receipt of a claim, an AI tool can suggest a document checklist, and an authorised officer can decide whether evidence meets policy requirements. The reviewer needs time, relevant source records and permission to reject the tool's output.

Block delegation when any gate fails

Confirm that the use of the data is permitted and the environment is approved. Test incorrect output, unsupported requests, changes in language or format, and interruption of service. For material decisions on people, payments, coverage or access, keep an authorised person responsible for the final action and give affected people a route to query or appeal it. Establish the controls with local risk, privacy, labour and legal advisers. NIST's generative-AI risk profile provides a reference for evaluation and monitoring [13].

Before any live use, ask a colleague to trigger the fallback from the normal workstation and document the result. If the team cannot reconstruct what happened, correct a record or resume service, narrow the scope or keep the work manual.

EQUIP / ALLOCATE

Train for the redesigned role and assign capacity

Training starts with a list of decisions the changed job requires. Have staff demonstrate routine use, source checking, handling an exception and completing the work when the tool is unavailable. A course certificate can record participation. A supervisor's observation on representative cases records performance.

Assessment	Practice and pass evidence
Check an AI draft	Compare every required field to the approved source. Correct an inserted error and log the correction.
Handle an exception	Identify a missing or disputed item. Route it to the person with authority and explain the decision in plain language.
Continue service	Complete a representative case with the tool switched off. Show where the record and handover go.

Set aside paid time and arrange cover. Offer accessible training formats, equipment and coaching across shifts. Give people who need more practice a supervised assignment and a chance to be reassessed. Preserve entry routes by planning mentored exposure to cases that new staff used to learn from. Use the ILO's quality-apprenticeship guidance as a design reference for structured learning and fair conditions [14]. Track progression and access separately from training attendance.

Put released hours on a roster

Measure staff minutes for a fixed window of eligible arrivals across all teams. Include prompting, verification, corrections, handovers and effort on cases that return to manual handling or remain open. Use the same follow-up period and case mix in both groups. Report unfinished cases and customer waiting time separately.

Weekly hours available = (baseline mean minutes per case − pilot mean minutes per case) × reference cases per week ÷ 60 − added weekly overhead hours.

Overhead covers recurring work outside case handling. Count fallback work once, under handling or overhead. Record one-off setup and training costs separately.

Assign positive net capacity to named work before expansion. Reserve it in the roster or backlog, track actual hours used, and ask the receiving team what improved. Reallocate partial hours only where the scheduling pattern makes them usable. If capacity is negative or has no recipient, revise the design or make an explicit service case for continuing. A positive time balance does not by itself justify a staffing decision.

VERIFY

Design the pilot before changing the workflow

Days	Decision and owner	What to record
1–15	Process owner and workers agree scope, baseline and service promise.	Eligible cases, case mix, exceptions, staff effort, queue age, outcome quality and worker feedback.
16–30	Risk lead, supervisor and training owner approve a shadow test.	Approved tools and data, performance on difficult cases, individual sign-off and a tested rollback.
31–60	Shift lead starts a bounded comparison with a named escalation route.	Matched case volumes, incidents, overrides, workload and the reasons cases leave the pilot.
61–90	Decision panel examines the evidence and hears worker feedback.	Who benefited, who absorbed new work, what happened to the time, and the scale, revise or stop decision.

In a shadow test, staff try the revised steps in parallel while the current process stays in charge. Start live testing only after staff, data and fallback checks pass. Move the decision date if the team needs more baseline or follow-up time. Observe at least two representative weeks and extend the baseline if volumes are low or the period is unusual. Agree the comparison before rollout. Where safe and feasible, randomly assign comparable cases or work periods. When staff share learning or queues across groups, assign teams or shifts and account for clustering. If you use matched shifts or a before–after comparison, record seasonal demand, policy or staffing changes and how those could explain the difference. Use established evaluation guidance when selecting the design [16].

Use separate case sets for training and testing. Have someone who did not build the workflow grade outputs against a fixed rubric, blinded to the workflow where feasible. Sample exceptions deliberately and report their results separately or weight them back to the actual mix for an overall estimate. Keep the same eligibility rules and track cases that were switched back to manual handling.

Collect only information needed for the pilot. Limit access to identifiable records, agree a retention schedule and report staff feedback in aggregates where this protects confidentiality. Invite concerns without requiring disclosure to a direct manager. Do not reuse a task ledger or AI logs to rank employees or infer an individual's replacement risk.

VERIFY / EVOLVE

Set thresholds and make the day-90 decision

Agree what would count as a meaningful benefit and a tolerable change before the pilot. Choose thresholds with workers, the service owner and risk. The examples below are decision rules for a bounded trial. Adapt them to consequences, baseline variation and sample size.

Measure	Practical definition	Example decision rule
Net staff capacity	Comparable weekly hours across all affected teams, after recurring overhead.	Set a local minimum, such as two hours a week for four comparable weeks.
Service quality	Cases that pass the same review criteria divided by cases reviewed. Also report counts of critical errors.	Set a maximum decline, such as two percentage points if your sample supports it. Pause after a serious control failure.
Customer access	Time to service and unresolved cases, including manual or assisted channels.	Investigate deterioration by channel, language and case type.
Worker load and skill	Overtime, rework and short worker check-ins, plus demonstrated skill among staff assigned to use the process.	Require sign-off for assigned staff. Check where extra work lands.
Use of capacity	Hours actually rostered and used for the agreed activity, alongside its outcome.	Review shortfalls with the receiving team before expanding.

Report denominator, missing cases and uncertainty next to each number. For timing, show the distribution and compare slow cases as well as typical ones. For quality, choose the number of reviews needed to support the maximum tolerable decline. Zero observed serious failures in a small trial gives limited information about rare harms. Keep close oversight or extend the trial if the evidence is too thin for a safe decision.

Scale within tested bounds when the local benefit clears the threshold, risks are controlled, people are ready and the service owner can show where capacity went. **Revise** when gains appear plausible but the workflow, training or data are incomplete. Time-box the revision and name the next test. **Stop or revert** when a hard gate fails, losses exceed benefit, or the team cannot demonstrate a safe operating boundary.

Publish the decision internally with worker feedback and any dissent. When conditions change, test again before extending authority. Trigger a review after a model, vendor, policy, data or staffing change, and set an ordinary review date even if nothing changes.

ILLUSTRATIVE CASE 01

Worked application in general insurance

Illustrative scenario. A Caribbean insurer wants to speed up motor claims intake. Its objective is to get complete files to assessors sooner and tell customers what evidence is outstanding. The trial covers a predefined queue of standard claims. A supervisor routes bodily injury, disputed coverage, suspected fraud and urgent exceptions to the established process.

Intake task	Proposed change	Check
Copy claim details into a second tracker	Retire if the authorised claims record meets all reporting and control needs.	Reconcile open cases from the source before removal.
Confirm receipt	Automate using an approved template and delivery log.	Verify customer and claim identifiers and provide an alternative channel.
Suggest missing documents	Assist with AI against the current approved checklist.	An officer checks each suggestion before contacting the customer.
Decide coverage, fraud referral or payment	Keep human-led.	A named officer records the grounds and follows the escalation policy.
Explain a complex case to a customer	Keep human-led and schedule time.	Review callbacks and sample whether customers understood their next step.

Calculate whether the team can use the hours

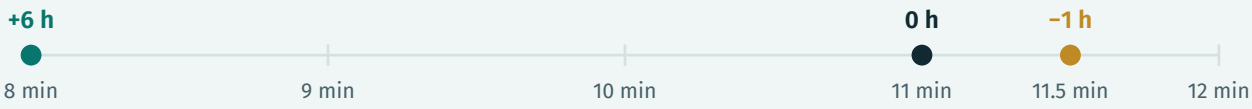
Assume 120 comparable eligible claims a week. Mean staff handling time falls from 12 to 8 minutes per case, including case-level checking and rework. That releases 8 hours before 2 additional hours of recurring team-level monitoring and fallback drills, leaving 6 hours a week. The supervisor reserves 3 hours for proactive callbacks, 2 for coached case reviews with new staff and 1 for complex files.

If assisted handling takes 11.5 minutes per case, the gain before overhead falls to 1 hour and the team loses 1 hour on balance. In that case, inspect correction time and handovers before deciding whether the service gained enough to justify the extra effort.

Check file completeness, repeat contacts and case age in both queues. Include any extra effort customers must make. The figures illustrate the calculation.

The trial breaks even at 11 minutes per claim

Illustrative net staff hours per week at different assisted handling times



ILLUSTRATIVE CASES 02–03

Worked applications in tourism and public services

A hotel reservations desk

Illustrative scenario. A small hotel wants to answer routine booking enquiries quickly while preserving a reliable route for guests with access needs. Remove duplicate entry where the approved booking system can supply the details. Use an AI draft for frequent questions, with staff checking rates, availability and terms against the booking system. Route accessibility arrangements, refunds and complaints to the duty lead. A rate or room type that the team cannot verify stays out of a response.

Teach reception staff to confirm a booking manually and practise the handover during a connectivity interruption. Test after-hours and multilingual enquiries before extending the hours of the service. Assess reply time alongside misquoted rates, repeat contacts and missed accessible arrangements. If staff can reclaim usable time, roster it for guest follow-up or training. If overnight review merely shifts work to the morning shift, record the transfer as a cost.

A public licensing office

Illustrative scenario. A licensing office wants residents to receive clearer guidance about incomplete applications. Staff compare the web and paper forms, remove duplicated fields only after the records owner agrees, and publish a single checklist. An approved system can send a standard receipt. A reviewer may use AI to draft a plain-language message based on an approved checklist, then confirm the required documents and sign it. An authorised officer decides eligibility and handles appeals under the applicable rules.

Compare completed applications and repeat visits by online and in-person channels. Include people who use assisted access or a paper form. Agree with staff how to process cases when the network fails. If a new workflow helps online applicants while the walk-in queue grows, review capacity and access before extending it.

Adapt to the local setting

Confirm languages, case volumes, service channels and the reliability of devices and connectivity before borrowing a pilot design. The World Bank's Eastern Caribbean digital transformation project identifies access and skills as regional implementation priorities [12]. The ILO's Caribbean platform-work study shows why an employee-only map may miss some people doing digital work [17]. Where contractors or shared services handle tasks, include their effort and the handovers they own in the assessment.

FIELD PROTOCOL

What would validate REWEAVE?

Primary question. Across participating Caribbean workplaces with comparable, low-risk workflows, does REWEAVE release more net staff time than a tool-first rollout while meeting predefined service quality, worker workload and safety thresholds?

The proposed contribution is an auditable combination: retire avoidable work before delegation, fund readiness for the changed role, and trace recovered time to a named activity before expansion. It builds on task-based economics, established job redesign and guidance for evaluating interventions [1–3, 16]. A literature-informed protocol and worked examples establish how to apply it. They cannot establish a causal effect or show that it transfers across organisations.

Proposed comparative test

Recruit workplaces that can follow the same measurement definitions and report a full baseline. Recruit across sectors and operating conditions, documenting which organisations decline or cannot participate. Compare REWEAVE with a tool-first pilot that has the same tool access, delivery period and mandatory safety controls. Add a current-workflow group where feasible to separate the effect of adopting a tool from the effect of redesigning work. Assign clusters such as teams where spillover is likely. If assignment is infeasible, document the comparison and treat causal interpretation as limited.

Before examining outcomes, register eligibility, primary and secondary measures, minimum worthwhile effects, quality margins, missing-data rules and stop conditions. Size the study for the primary contrast and the quality margin, allowing for differences between teams and incomplete observations. Use all assigned clusters in the main analysis where feasible, report changes by original assignment, and show uncertainty and protocol deviations. Capture actual fidelity to each stage so a null result can be interpreted.

Primary estimate. Difference between groups in change from baseline in net staff hours per 100 eligible cases, adjusted for case mix. Report service quality, workload, demonstrated skill and where capacity went beside the primary estimate. Examine variation by site and case type only with enough observations. Account for setup, tool and training costs separately from recurring hours.

Evidence that would change the decision

A time gain fails the joint decision rule when quality, workload or access breaches the agreed limits. A serious control failure triggers the stop rule regardless of the averages. No meaningful difference after adequate measurement would weaken the case for the framework's extra steps. A small or inconclusive trial would call for a better test. To assess repeatability, ask two reviewers to classify a sample of tasks independently and compare agreement, particularly on cases where one reviewer delegates and the other keeps a human lead.

> **Evidence status** Conceptual framework and proposed field protocol. No participant records or field results accompany this edition.

WORKSHEETS

Worksheet 1 · One card per task

Copy these fields into a shared document. Complete one card with the person who performs the task and the person who receives the output. Bring a routine case and an exception.

Field	What to write down
Workflow, task and date	Give the task a verb and an object. Name the service outcome, local site and version.
Trigger, output and receiver	Say who starts it, what it produces and who uses the result.
Volume and effort	Record the time window, eligible case count and staff minutes. Separate waiting from handling.
Exceptions and dependencies	List difficult cases, channels, systems, downstream teams and any control this task supports.
Learning or judgement	Describe the decision the task teaches, the discretion needed and what a trainee should practise.
Task decision	Retire, simplify or keep as is. For a task that remains, choose human-led, AI-assisted or automated within limits. Record the reason and the alternative considered.
Accountable person	Name the person who can sign off, review exceptions and order a return to the prior process.
Evidence and data boundary	Identify the approved source, permitted inputs, who may see them and the quality checks.
Fallback	State what triggers it, who acts, where the manual instructions live and when the team last tested it.
Open question and next date	Record what remains unknown, who will find out and when the decision comes back for review.

Check before you leave the session

Ask the receiver what would break if the task disappeared. Record the answer and who will check it before the change.

Decision record Task ID _____ Owner _____ Review date _____

A first session for an individual worker

List your repeated tasks, one task that teaches you judgement and one you postpone when the queue is full. Test a possible shortcut on public or dummy material within your permissions. Bring the checking time, failed cases and proposed use of any hours saved to your supervisor before changing a shared process.

WORKSHEETS

Worksheet 2 · Pilot and release decision

Copy the prompts into a shared decision record before the pilot.

Before live use	Record your agreement
Workflow, scope and owner	Which cases qualify? Which cases use the established process? Who can pause the pilot?
Baseline and comparison	What will you compare? Which dates, case types and teams are included? Who will verify the case counts?
Safety and service thresholds	What event stops the pilot? How much quality or access deterioration is tolerable? What evidence will support that limit?
Staff readiness	Who needs paid practice and cover? Who will sign off capability on routine work, exceptions and fallback?
Time budget	What reference volume will you use? Which case-level and recurring tasks will you count? Where will released hours go?
Review	Who hears worker feedback and who decides? When will you revisit the design after a change?

At the decision meeting	Record evidence and outcome
Service	Quality review count, critical events, repeat contacts, slow cases and channel differences.
Capacity and workload	Net staff time, overtime or work transferred, and actual hours used in the destination activity.
Readiness and access	Staff demonstrations, missing training access and changes in entry-level learning opportunities.
Decision and evidence gaps	Scale, revise or stop. List which groups and scenarios remain outside the decision.
Signature and reopening trigger	Process owner, worker representative or participant delegate, and risk reviewer where relevant. Name the date and trigger.

Record the redesigned job

Write the role's purpose, what work it gains or gives up, who can make each material decision and what changes on the roster. Record paid practice, the route for new hires to learn the job, and any option for reassignment or progression. Capture worker feedback and disagreements, name the role owner and date the next review.

EVIDENCE REGISTER

References

[1] Acemoglu, D., & Restrepo, P. (2019). *Automation and new tasks: How technology displaces and reinstates labor*. Journal of Economic Perspectives, 33(2), 3–30. doi.org/10.1257/jep.33.2.3

[2] Jesuthasan, R., & Boudreau, J. W. (2022). *Work without jobs: How to reboot your organization's work operating system*. MIT Press. doi.org/10.7551/mitpress/13930.001.0001

[3] Morgeson, F. P., & Humphrey, S. E. (2006). *The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work*. Journal of Applied Psychology, 91(6), 1321–1339. doi.org/10.1037/0021-9010.91.6.1321

[4] Gmyrek, P., Berg, J., Kamiński, K., Konopczyński, F., Ładna, A., Nafradi, B., Rosłaniec, K., & Troszyński, M. (2025). *Generative AI and jobs: A refined global index of occupational exposure*. ILO Working Paper 140. DOI 10.54394/HETP0387. ILO publication and full text

[5] Gmyrek, P., Winkler, H., & Garganta, S. (2024). *Buffer or bottleneck? Employment exposure to generative AI and the digital divide in Latin America*. ILO Working Paper 121, joint ILO–World Bank publication. DOI 10.54394/TFZY7681. ILO publication and full text

[6] Brynjolfsson, E., Li, D., & Raymond, L. R. (2025). *Generative AI at work*. Quarterly Journal of Economics, 140(2), 889–942. doi.org/10.1093/qje/qjae044

[7] Dell'Acqua, F., et al. (2026). *Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality*. Organization Science, 37(2), 403–423. doi.org/10.1287/orsc.2025.21838

[8] Becker, J., Rush, N., Barnes, E., & Rein, D. (2025). *Measuring the impact of early-2025 AI on experienced open-source developer productivity*. Preprint. doi.org/10.48550/arXiv.2507.09089

[9] METR. (2026, 24 February). *We are changing our developer productivity experiment design*. Research update. METR

EVIDENCE REGISTER

References continued

[10] Lane, M., Williams, M., & Broecke, S. (2023). *The impact of AI on the workplace: Main findings from the OECD AI surveys of employers and workers*. OECD Social, Employment and Migration Working Papers, No. 288. doi.org/10.1787/ea0a0fe1-en

[11] UNESCO. (2024). *Caribbean Artificial Intelligence Policy Roadmap*. UNESCO

[12] World Bank. (2020, 22 June). *First-time financing by World Bank for digital economy in the Eastern Caribbean approved for US$94 million*. Caribbean Digital Transformation Project announcement. World Bank

[13] Autio, C., et al. (2024). *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*. NIST AI 600-1. doi.org/10.6028/NIST.AI.600-1

[14] International Labour Organization. (2023). *Quality Apprenticeships Recommendation, 2023 (No. 208)*. ILO Recommendation No. 208

[15] Noy, S., & Zhang, W. (2023). *Experimental evidence on the productivity effects of generative artificial intelligence*. Science, 381(6654), 187–192. doi.org/10.1126/science.adh2586

[16] HM Treasury. (2026). *Magenta Book: Central Government guidance on evaluation*. GOV.UK

[17] International Labour Organization. (2025). *Decent work in the platform economy in the Caribbean*. International Labour Organization

Suggested citation Dunkley, A. (2026). *REWEAVE: Reassessing jobs in the age of AI*. Caribbean AI Association. Consultation edition v0.1.

> **Publication status** This consultation edition proposes a method for testing. The three case studies use illustrative assumptions. External peer review and workplace validation remain pending.