FJB

JOSEFF BETANCOURT

AI PLATFORM ARCHITECT | JOSEFFB.COM

AI-Assisted Estimation Methodology

A reusable WBS, calibration, and stage-gate method

Separate baseline effort, AI-assisted active effort, operator load, and elapsed schedule.

APPLIED METHODOLOGY WORKING PAPER

Revised 13 August 2026

Version 1.2.0

Author and publisher: Joseff Betancourt

Publication version

Contents

Document Metadata

Field	Value
Canonical title	*AI-Assisted Estimation Methodology*
Version	1.2.0
Document type	Applied methodology and evidence-informed working paper
Status	Publication version
Revision date	13 Aug. 2026
Bounded review method	Narrative audit of nine primary or authoritative sources, including the exact early-2025 METR study. Claims were checked against the issuing organization, original paper, or current official handbook page. This was not a systematic review or meta-analysis.
Empirical status	The methodology has not been independently validated. External study outcomes are presented as source-reported findings. The 10:1, 40:1, and 15/75/10 planning inputs have no supporting stage-level dataset in this publication version and remain hypotheses requiring local calibration.
DOI	Not assigned
Canonical URL	Not assigned
License	Pending author decision
Author ORCID	Not assigned
Suggested citation (MLA 9)	Betancourt, Joseff. *AI-Assisted Estimation Methodology*. Version 1.2.0, 13 Aug. 2026. Applied methodology working paper.

Abstract

AI-assisted delivery redistributes engineering effort across human decisions, human-AI supervision, machine execution, verification, and acceptance. This paper presents a work-breakdown-structure method for estimating that blended active effort without treating model runtime as total project effort. It defines separate estimation objects, classifies work by delivery pattern, applies explicit planning and acceptance floors, calculates low/high scenarios, and requires calibration against completed work. A bounded review of controlled experiments, field experiments, mature-repository research, and authoritative estimating and risk-management guidance finds no reviewed basis for one universal AI multiplier. The included 10:1, 40:1, and 15/75/10 values are therefore reference hypotheses, not empirical benchmarks or commitment rates.

Keywords: AI-assisted estimation; software cost estimation; work breakdown structure; human-AI collaboration; developer productivity; calibration; stage-gate delivery; sensitivity analysis; active effort; empirical uncertainty

Epistemic Status Protocol

The labels below apply to material claims in this publication:

Label	Meaning in this document
FACT	A document definition, verified source identity, or reproducible arithmetic relationship. It does not convert a cited study's outcome into a universal fact.
REPORTED	A result, requirement, or conclusion stated by a cited source and not independently replicated for this paper.
INFERENCE	A bounded synthesis drawn from the reviewed sources or from the method's structure.

Label	Meaning in this document
HYPOTHESIS	A testable but unvalidated local assumption, coefficient, allocation, or expected directional effect.
OPINION	A normative methodological or governance choice recommended by the author.

Executive Summary

REPORTED — Estimating foundation. The U.S. Government Accountability Office's cost guide treats a work breakdown structure (WBS), explicit assumptions, sensitivity and risk analysis, documented results, and updates with actuals as parts of reliable estimating. NASA's current Software Engineering Handbook Version D likewise uses the WBS as an estimating and planning framework and requires documented software cost-estimation models for covered project classes (United States Government Accountability Office; National Aeronautics and Space Administration, “Work Breakdown Structures”; National Aeronautics and Space Administration, “Cost Estimation”).

AI-assisted delivery changes how engineering effort is distributed, but it does not eliminate planning, review, acceptance, security judgment, coordination, or deployment work. A credible estimate must therefore measure more than model runtime or code-generation speed.

This methodology starts with a conventional WBS and a traditional effort baseline. Each work package is then classified by its delivery pattern:

Work pattern	Responsible role	Planning coefficient
Human-only decision or acceptance	Human	1:1
Planning, clarification, and supervisory iteration	Human/AI iterative loop	Configurable; 10:1 only when calibrated
Bounded implementation and AI QA	AI	Configurable; 40:1 only when calibrated

HYPOTHESIS — Reference planning profile. For scenario analysis, an AI-led stage may start with this allocation:

- 15% Human/AI planning and clarification.
- 75% AI implementation and AI QA.
- 10% human review and acceptance.

The percentages and ratios are planning inputs, not findings established by the reviewed research. Teams must calibrate them against comparable completed work.

FACT — Estimation object. The output is an **AI-assisted active-effort estimate** as defined in this paper. It is not model runtime, traditional person-hours, staffing capacity, cost, or elapsed calendar time.

INFERENCE — Research position. The reviewed studies report materially different outcomes across a bounded coding exercise, enterprise field experiments, and mature-repository work. Within this evidence set, task class, workflow conditions, measurement design, and human review are more defensible inputs than a single universal multiplier (Peng et al.; Cui et al.; Becker et al., “Measuring”; DORA).

1. Purpose

Use this methodology to:

- Create a transparent WBS for AI-assisted delivery.
- Identify which work is AI-compressible and which work remains human-bound.
- Compare traditional and AI-assisted effort without hiding oversight.
- Produce low/high ranges and sensitivity scenarios.
- Convert an effort estimate into a schedule only after capacity and dependencies are known.
- Replace assumptions with actual project evidence over time.

The method is suitable for software development, technical documentation, integration projects, data work, operational automation, and other projects where AI produces bounded artifacts that humans direct and accept.

2. What the Reviewed Evidence Supports

REPORTED — External productivity evidence. These outcomes belong to their specific research settings:

Evidence	Setting	Source-reported result	Bounded methodological lesson
GitHub Copilot controlled experiment	95 Upwork-recruited programmers completing a bounded JavaScript HTTP-server task	The treatment group completed the task 55.8% faster than the control group	A bounded greenfield task can show substantial acceleration; it does not represent all repository work (Peng et al.)
Microsoft, Accenture, and Fortune 100 field experiments	Three randomized field experiments covering 4,867 developers using code-completion assistance	The pooled analysis reported a 26.08% increase in completed tasks, with a 10.3% standard error; individual experiments were noisy	Enterprise effects can be positive while remaining heterogeneous and measurement-dependent (Cui et al.)
Exact early-2025 METR RCT	16 experienced open-source developers, 246 tasks, mature repositories, and an average of about five years' prior repository experience	Allowing February-June 2025 AI tools increased completion time by 19% in this setting	Familiar, mature repositories and high quality/context demands can produce a result opposite to bounded greenfield findings (Becker et al., “Measuring”)
METR 2026 experiment-design update	A later study with 10 returning and 47 newly recruited developers	METR judged the later central estimate unreliable because of participation and task-selection effects, the lower pay rate, and time-accounting problems when developers ran agents concurrently	Agentic workflows require measures that separate operator time, agent time, parallelism, selection, and task mix (Becker et al., “Changing”)
DORA 2025	Organization-level research on AI-assisted software delivery	DORA characterized AI as an amplifier of existing organizational strengths and weaknesses	Tool access alone does not determine system-level delivery performance (DORA)

INFERENCE — Bounded conclusion. These reviewed sources do not establish a universal AI productivity multiplier. They examine different populations, tasks, tools, outcomes, and time periods, so their reported effects should not be pooled into one coefficient without a justified comparison model.

REPORTED — Estimating and governance evidence.

Source	What the source supports in this methodology
GAO Cost Estimating and Assessment Guide	Define purpose and scope, build a WBS, document assumptions, select methods, analyze sensitivity and risk, present results, and update the estimate with actuals (United States Government Accountability Office)
NASA Software Engineering Handbook Version D, Topic 7.05	Use a WBS to define and iteratively refine the complete work needed for planning, estimating, organization, and control (National Aeronautics and Space Administration, “Work Breakdown Structures”)
NASA Software Engineering Handbook Version D, SWE-015	Establish, document, maintain, and update software cost-estimation models and parameters appropriate to the project class and uncertainty (National Aeronautics and Space Administration, “Cost Estimation”)
NIST AI RMF 1.0	Document context, human oversight, methods, metrics, uncertainty, testing, and limits on generalization; keep risk measurement and management iterative (National Institute of Standards and Technology 25, 28-30)

HYPOTHESIS — Local reference profile. The 10:1 Human/AI and 40:1 AI coefficients in the worked profile originated as operator-reported workflow observations. This publication version contains no task-level sample, measurement protocol, dispersion, or replication that validates them. Treat them as scenario coefficients until comparable stage actuals support another status.

OPINION — Conservative planning rule. Until stage actuals exist, retain the traditional estimate as the commitment baseline and show AI-assisted values as explicit scenarios.

3. Estimation Objects

FACT — Document definitions. This methodology keeps the following quantities separate:

Quantity	Definition	Appropriate use
Traditional baseline effort	Estimated effort for the same accepted outcome under the comparison workflow	Scope baseline and historical comparison
AI-assisted active effort	Active work consumed by the blended Human, Human/AI, and AI workflow	Delivery planning and option comparison
Operator-active effort	Time a human actively scopes, directs, reviews, corrects, decides, or accepts	Staffing and workload planning
AI execution time	Time agents actively execute, including unattended runs	Tool capacity and concurrency planning
Elapsed schedule	Calendar duration after dependencies, capacity, queues, and change windows	Delivery commitments
Cost	Labor, platform, model, infrastructure, and contingency costs	Budgeting

Never present AI execution time as total project effort. Never present active effort as elapsed duration.

4. Work Classification

OPINION — Classification rule. Classify work by the accountable delivery pattern and the evidence needed for acceptance, not by whether an AI tool can produce a plausible artifact.

4.1 Human

Use Human at 1:1 when the outcome requires accountable human action:

- Stakeholder interviews.
- Final scope or architecture decisions.
- Risk acceptance.
- Business and user acceptance.
- Security or legal approval.
- Production authorization.
- Organizational coordination.

AI may assist with preparation, but it does not remove the accountable action.

4.2 Human/AI Iterative Loop

Use Human/AI for supervisory work where the operator repeatedly provides context, asks questions, reviews alternatives, and corrects direction:

- Discovery synthesis.
- Requirements clarification.
- Architecture and design.
- Stage planning.
- Test and acceptance planning.
- Investigation of ambiguous legacy behavior.
- Security-sensitive review.
- Cross-system integration decisions.

The loop ratio must include operator interaction. It is not a count of the minutes required for a model to return an answer.

4.3 AI

Use AI for bounded execution that can be delegated to a proven workflow:

- Scaffolding and implementation.
- Mechanical refactoring.
- Test generation and execution.
- Documentation drafting.
- Data transformation.
- Repeatable analysis.
- Evidence assembly.
- AI-led QA against objective gates.

AI classification requires clear context, an authorized mutation boundary, testable acceptance criteria, and a human owner.

5. Stage-Gate Delivery Pattern

Even when a project begins with a large plan, major stages normally require a new clarification loop. Unknowns surface after implementation begins, and operators pause to answer questions between phases.

HYPOTHESIS — Reusable starting allocation. An AI-led stage may begin with:

Component	Default share	Role	Reference ratio
Plan and clarify	15%	Human/AI iterative loop	10:1
Implement and self-test	75%	AI	40:1
Review and accept	10%	Human	1:1

The shares and ratios are scenario inputs. Change them when project evidence shows a different operating pattern.

Planning-only and acceptance-only WBS rows should be represented separately instead of forcing them through the blended-stage formula.

6. Core Calculation

For WBS item `i`:

- `B_i` = traditional baseline hours.
- `P_i` = planning share.
- `A_i` = AI-execution share.
- `H_i` = human-acceptance share.
- `R_loop` = Human/AI-loop ratio.
- `R_ai` = AI-execution ratio.
- `R_human` = 1.
- `P_i + A_i + H_i = 1`.

The AI-assisted active effort is:

```
E_i = B_i * ((P_i / R_loop) + (A_i / R_ai) + (H_i / R_human))
```

Using the reference-stage hypothesis:

```
E_i = B_i * ((0.15 / 10) + (0.75 / 40) + (0.10 / 1))
E_i = B_i * 0.13375
```

FACT — Conditional arithmetic. If those shares and ratios are accepted, `1 / 0.13375 = 7.4766...`; the implied blended compression is approximately 7.5:1, not 40:1, because planning and acceptance remain in the calculation.

This arithmetic does not validate its inputs. The result is usable only to the extent that the selected coefficients match the team, task class, comparison baseline, and acceptance standard. Otherwise use locally calibrated values or a conservative 1:1 baseline.

7. Calibration Ladder

Level	Evidence	Permitted use
U0 - Uncalibrated	No comparable completed work	Keep 1:1 as the commitment baseline; show AI scenarios separately
U1 - Pilot	A small number of representative completed tasks	Use a broad low/high range and label it preliminary
U2 - Workflow-calibrated	Repeated work with the same operator, tools, controls, and task class	Use observed medians and dispersion for internal ROM planning
U3 - Proven orchestrated workflow	Stable context preparation, execution, testing, review, correction, and evidence practices across multiple stages	Use higher compression only for matched work classes
U4 - Continuously calibrated	Actuals collected and ratios updated by task class	Use for recurring forecasts with documented confidence

OPINION — Transfer rule. A high ratio from one project should not be reused as a general rate. Transfer requires a documented match in task class, operator practice, tools, controls, baseline definition, and acceptance threshold.

8. Estimation Procedure

Step 1: Define the accepted outcome

Document:

- Included deliverables.
- Explicit exclusions.
- Environments.
- Quality and security requirements.
- Acceptance owner.
- Deployment and support responsibilities.

Step 2: Build the traditional WBS

Include the whole outcome, not only code generation:

- Planning and monitoring.
- Requirements and design.
- Implementation.
- Testing and repair.
- Security and compliance.
- Integration.
- Documentation and training.
- Stakeholder review.
- Deployment and handoff.

This completeness rule aligns with the GAO estimating process and NASA's software WBS guidance (United States Government Accountability Office; National Aeronautics and Space Administration, “Work Breakdown Structures”).

Step 3: Create the traditional baseline

Estimate each WBS row as a low/high range in hours. Record the source:

- Historical actual.
- Analogy.
- Bottom-up engineering judgment.

- Model-based estimate.
- Expert judgment.

Do not derive the traditional baseline by reversing a desired AI result.

Step 4: Classify each WBS row

Select Human, Human/AI iterative loop, AI, or an explicitly blended stage. Record why the work is suitable for that classification.

Step 5: Set coefficients

Use local actuals when available. When actuals are unavailable:

- Keep the traditional estimate as the commitment baseline.
- Present AI coefficients as scenarios.
- Identify what must be proven before using the optimistic scenario.

Step 6: Add human floors

Add explicit effort for:

- Inter-stage planning and clarification.
- Architecture and security judgment.
- Operator review and correction.
- UAT and acceptance.
- Production authorization.
- Required documentation and evidence.

Step 7: Calculate low/high AI-assisted effort

Apply the formula row by row. Sum unrounded values and round only the displayed totals.

Step 8: Test sensitivity

At minimum, vary:

- AI ratio.
- Human/AI-loop ratio.
- Planning share.
- Acceptance share.
- Rework rate.
- Percentage of work that is actually AI-suitable.

Step 9: Convert effort to schedule

Only after the effort estimate is complete, apply:

- Available operator hours.
- Agent concurrency.
- Dependency sequence.
- Review capacity.
- Environment access.
- Approval and deployment windows.

Vacations, holidays, weekends, leave, queues, and waiting periods are excluded from active-effort hours but must be included in elapsed schedule planning.

Step 10: Update with actuals

At every stage gate, record:

- Estimated and actual active effort.
- Operator-active time.
- AI execution time.
- Rework and failed-run time.
- Acceptance defects.
- Causes of variance.
- Whether the task classification was correct.

NASA's Version D cost-estimation guidance calls for estimates and their parameters to be maintained as the project evolves, while NIST calls for context-connected metrics, documented uncertainty, and ongoing measurement (National Aeronautics and Space Administration, “Cost Estimation”; National Institute of Standards and Technology 28-30).

9. Risk and Adjustment Factors

HYPOTHESIS — Directional adjustment factors. The factors below are expected to affect achievable compression, but their magnitude and sometimes their direction must be measured locally.

Factors that may reduce achievable compression:

- Mature or poorly documented legacy code.
- Weak automated tests.
- High coupling across systems.
- Security, privacy, legal, or regulatory sensitivity.
- Unstable requirements.
- Restricted environments or credentials.
- Novel frameworks or domains.
- Large unstructured context.
- Low operator familiarity with the codebase.
- Unproven orchestration or weak review practices.

Factors that may improve compression:

- Small, bounded work packages.
- Strong tests and deterministic acceptance gates.
- Stable architecture and coding conventions.
- High-quality context packs.
- Reusable automation.
- Fast environments and reliable tooling.
- Experienced operators.
- Repeated work in the same task class.

Risk adjustments must change a named assumption or coefficient. An unexplained blanket contingency is weaker than a traceable sensitivity scenario.

10. Basis of Estimate

Every estimate issued under this methodology should state:

Field	Required content
Scope baseline	Included outcome and exclusions
WBS version	Identifier and date
Traditional basis	Historical, analogy, bottom-up, model, or judgment
Role classification	Human, Human/AI, AI, or blended
Coefficients	Ratios and evidence level
Stage shares	Planning, AI execution, and acceptance percentages
Human floors	Decisions, reviews, UAT, approvals, and deployment
Operator assumptions	Experience, availability, and codebase familiarity
Tool assumptions	Models, agents, environments, and concurrency
Risk range	Low/high cases and sensitive drivers
Schedule conversion	Capacity, dependencies, and calendar assumptions
Validation plan	How actuals will be captured and ratios updated

11. Worked Example

HYPOTHESIS — Scenario inputs. Assume an AI-led implementation stage has a traditional baseline of 80-120 hours and, solely for this example, the team accepts the 15/75/10 allocation, 10:1 Human/AI ratio, and 40:1 AI ratio.

Component	Baseline share	Low baseline	High baseline	Ratio	AI-assisted low	AI-assisted high
Planning loop	15%	12	18	10:1	1.2	1.8
AI implementation	75%	60	90	40:1	1.5	2.25
Human acceptance	10%	8	12	1:1	8.0	12.0
Total	**100%**	**80**	**120**	**Blended**	**10.7**	**16.1**

FACT — Scenario arithmetic. The unrounded totals are 10.7 and 16.05 hours; the displayed high total rounds to 16.1 hours.

INFERENCE — Scenario behavior. Under these inputs, human acceptance dominates the blended result even though the AI-execution component is assigned the largest baseline share. This is a property of the formula and chosen inputs, not proof that human acceptance dominates every AI-assisted workflow.

12. Presentation Standard

Present leadership with:

1. Traditional low/high effort.
2. AI-assisted low/high active effort.
3. Operator-active load.
4. Expected elapsed schedule.
5. Primary assumptions and exclusions.
6. Confidence/calibration level.
7. Top sensitivity drivers.
8. Decision: GO, SPLIT, or NO-GO / WAIT when capacity is known.

Do not present a multiplier without its task class, comparison baseline, evidence status, and human floors.

13. Quality Controls

- The estimator and acceptance owner should be identifiable.
- AI must not approve its own security, legal, or business-risk decisions.
- WBS rows must have testable completion evidence.
- Ratios must be traceable to a scenario or actuals.
- Estimates must be versioned.
- Material scope changes require re-estimation.
- The estimate must distinguish observed actuals from proposed coefficients.
- Model or tool upgrades do not automatically justify a new ratio.
- Parallel agents may reduce elapsed duration without reducing total oversight.
- Actuals must include failed runs, correction, and rework.
- Generalizability limits must be documented when evidence comes from a different task or deployment context, consistent with NIST's measurement guidance (National Institute of Standards and Technology 28-30).

14. Reusable WBS Schema

Column	Description
WBS ID	Stable identifier
Phase	Major delivery phase
Work package	Accepted outcome
Baseline low/high	Traditional effort hours
Work pattern	Human, Human/AI, AI, or blended
Planning/AI/Human shares	Sum to 100%
Loop/AI/Human ratios	Evidence-backed coefficients
AI-assisted low/high	Formula result
Confidence	Calibration level and uncertainty
Dependencies	Inputs and sequencing
Acceptance evidence	Objective completion proof
Actual effort	Recorded after completion
Variance notes	Why actual differed

The companion `AI-Assisted-Estimation-Worksheet.csv` implements this schema.

15. Adoption Checklist

[ ] Define scope and acceptance owner.

[ ] Build a complete traditional WBS.

[ ] Record traditional low/high effort and basis.

[ ] Classify every row.

[ ] Add inter-stage planning and human acceptance.

[ ] Select coefficients and evidence level.

[ ] Run sensitivity scenarios.

[ ] Separate effort, capacity, schedule, and cost.

[ ] Review assumptions with delivery and acceptance owners.

[ ] Version and publish the basis of estimate.

[ ] Capture actuals at each stage gate.

[ ] Recalibrate before reusing ratios.

16. Research Interpretation

REPORTED. The reviewed authoritative guidance supports WBS discipline, explicit assumptions, documented parameters, sensitivity analysis, human oversight, continuous measurement, and limits on generalization (United States Government Accountability Office; National Aeronautics and Space Administration, “Cost Estimation”; National Institute of Standards and Technology 25, 28-30).

REPORTED. The reviewed productivity studies report heterogeneous effects: faster completion in a bounded greenfield task, more completed tasks in pooled enterprise field experiments, and slower completion in METR's early-2025 mature-repository setting (Peng et al.; Cui et al.; Becker et al., “Measuring”).

INFERENCE. The comparison supports calibration by task class and workflow more strongly than it supports transferring any one study's effect size into a planning coefficient.

HYPOTHESIS. The 10:1, 40:1, and 15/75/10 reference profile may be useful for scenario analysis in an orchestrated workflow with explicit human floors. Its accuracy remains unknown until stage-level actuals are captured.

OPINION. A decision-useful estimate should expose that uncertainty instead of presenting the most optimistic AI ratio as a commitment.

17. Review Boundaries and Limitations

- This was a bounded narrative review, not a systematic literature review, meta-analysis, or exhaustive search of AI productivity research.
- Sources were included because they directly support the methodology's existing estimating, governance, or productivity claims. Absence from this review is not evidence against another study.
- Source-reported outcomes were checked against primary papers or issuing organizations, but their data and analyses were not independently replicated.
- The studies use different tasks, populations, tools, quality thresholds, periods, and outcome measures. Direct numerical comparison is descriptive, not a pooled causal estimate.
- METR explicitly warns against generalizing its early-2025 RCT to most developers or other domains, and its 2026 update describes selection and time-accounting problems in the later experiment (Becker et al., “Measuring”; Becker et al., “Changing”).
- DORA's organization-level amplifier framing is not a task-level compression coefficient (DORA).
- The local planning coefficients have no documented sample size, actuals table, confidence interval, or independent validation in this publication version.
- Model capability, tool design, adoption, and workflow practice change over time. Any future publication should recheck time-sensitive evidence and URLs.
- This methodology estimates active effort. It does not by itself establish staffing capacity, elapsed schedule, cost, safety, legal compliance, or production readiness.

Works Cited

Becker, Joel, et al. “Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity.” *Model Evaluation & Threat Research*, 10 July 2025, https://metr.org/Early_2025_AI_Experienced_OS_Devs_Study-paper.pdf. Accessed 13 Aug. 2026.

Becker, Joel, et al. “We Are Changing Our Developer Productivity Experiment Design.” *Model Evaluation & Threat Research*, 24 Feb. 2026, https://metr.org/blog/2026-02-24-uplift-update/. Accessed 13 Aug. 2026.

Cui, Zheyuan (Kevin), et al. “The Effects of Generative AI on High-Skilled Work: Evidence from Three Field Experiments with Software Developers.” *Microsoft Research*, June 2025, https://www.microsoft.com/en-us/research/publication/the-effects-of-generative-ai-on-high-skilled-work-evidence-from-three-field-experiments-with-software-developers/. Accessed 13 Aug. 2026.

DORA. *State of AI-Assisted Software Development 2025*. Google Cloud, 2025, https://dora.dev/research/2025/dora-report/. Accessed 13 Aug. 2026.

National Aeronautics and Space Administration. “7.05 - Work Breakdown Structures That Include Software.” *NASA Software Engineering Handbook, Version D*, https://swehb.nasa.gov/spaces/SWEHBVD/pages/102695623/7.05+-+Work+Breakdown+Structures+That+Include+Software. Accessed 13 Aug. 2026.

National Aeronautics and Space Administration. “SWE-015 - Cost Estimation.” *NASA Software Engineering Handbook, Version D*, https://swehb.nasa.gov/spaces/SWEHBVD/pages/102695400/SWE-015+-+Cost+Estimation. Accessed 13 Aug. 2026.

National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. U.S. Department of Commerce, Jan. 2023, NIST AI 100-1, https://doi.org/10.6028/NIST.AI.100-1.

Peng, Sida, et al. “The Impact of AI on Developer Productivity: Evidence from GitHub Copilot.” *arXiv*, 13 Feb. 2023, arXiv:2302.06590, https://doi.org/10.48550/arXiv.2302.06590.

United States Government Accountability Office. *Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Program Costs*. GAO-20-195G, Mar. 2020, https://www.gao.gov/products/gao-20-195g. Accessed 13 Aug. 2026.

Publication and Attribution

Publication field	Information
Version	1.2.0
Status	Publication version
Original prepared date	3 Aug. 2026
Current revision date	13 Aug. 2026
Author and publisher	Joseff Betancourt
Website	joseffb.com
Research support	Fernain Research

Research direction and methodology: Joseff Betancourt. Evidence discovery, analysis, adversarial review, and drafting: AI-assisted. Final evaluation, editorial judgment, and publication responsibility: Joseff Betancourt.