

# AI and Machine Learning Data Quality Checklist - Plan a Small Project

AI and Machine Learning | Plan a Small Project | Asset ID TJ-12-04-03

Generated or predicted output can be wrong or biased. Protect private data and require human review for legal, medical, financial, safety or other consequential uses.

## Purpose and working method

This independent, printable companion helps you connect model objectives, data, evaluation and human oversight to a useful decision while working toward one outcome: turn a useful idea into owned steps and acceptance evidence. Use a suitable calculator or utility for computation, then use these pages to preserve meaning, assumptions, checks and decisions. A number alone is not a decision record; a useful record shows where the inputs came from, why the method fits and what would make the result unsafe or misleading.

| Project snapshot     | Record                              |
|----------------------|-------------------------------------|
| Project or question  | _____                               |
| Prepared by / date   | _____                               |
| Decision owner       | _____                               |
| Intended outcome     | Plan a Small Project                |
| Tool or method used  | Name and version/date: _____        |
| Authoritative source | Title, owner and access date: _____ |
| Review deadline      | _____                               |

## Domain readiness map

| # | Input area            | Current state               | Trace         |
|---|-----------------------|-----------------------------|---------------|
| 1 | User and task         | Known / estimated / missing | Source: _____ |
| 2 | Data and consent      | Known / estimated / missing | Source: _____ |
| 3 | Baseline method       | Known / estimated / missing | Source: _____ |
| 4 | Model assumptions     | Known / estimated / missing | Source: _____ |
| 5 | Evaluation measures   | Known / estimated / missing | Source: _____ |
| 6 | Failure and oversight | Known / estimated / missing | Source: _____ |

## Data Quality Checklist

This working page applies the data quality checklist format to plan a small project. Complete it before polishing a conclusion. The useful evidence is scope, owner, dependencies, estimate and done condition; the governing rule is: start only when the first milestone and stop condition are clear.

| Quality check         | Method                                         | Finding | Action |
|-----------------------|------------------------------------------------|---------|--------|
| User and task         | test whether the data is fit for this decision | _____   | _____  |
| Data and consent      | test whether the data is fit for this decision | _____   | _____  |
| Baseline method       | test whether the data is fit for this decision | _____   | _____  |
| Model assumptions     | test whether the data is fit for this decision | _____   | _____  |
| Evaluation measures   | test whether the data is fit for this decision | _____   | _____  |
| Failure and oversight | test whether the data is fit for this decision | _____   | _____  |

### Completion standard

- Write real units, dates and source owners; avoid labels such as normal, typical or standard without a definition.
- Mark estimates as estimates. Record a range when a single value would imply false precision.
- Keep an untouched copy of source data and give every working version a date or revision.
- Record contrary evidence. A workbook that only supports the preferred answer is not a reliable check.
- Remove personal, confidential or security-sensitive details that are unnecessary for the decision.

## Input and Data Map

**Privacy gate: minimise personal, confidential and security-sensitive information. If a value can identify a person, account, location or protected system, use an approved environment and retention rule.**

| # | Data element          | Minimum validation       | Source       | Status      |
|---|-----------------------|--------------------------|--------------|-------------|
| 1 | User and task         | identity and date        | Owner: _____ | Ready / gap |
| 2 | Data and consent      | definition and unit      | Owner: _____ | Ready / gap |
| 3 | Baseline method       | range and precision      | Owner: _____ | Ready / gap |
| 4 | Model assumptions     | missing-value rule       | Owner: _____ | Ready / gap |
| 5 | Evaluation measures   | consistency with peers   | Owner: _____ | Ready / gap |
| 6 | Failure and oversight | fitness for the decision | Owner: _____ | Ready / gap |

## Transformation trail

| Stage       | Question                                       | Record                    |
|-------------|------------------------------------------------|---------------------------|
| Raw input   | What arrived before cleaning or conversion?    | Location / version: _____ |
| Cleaning    | What was removed, repaired or standardised?    | Rule: _____               |
| Conversion  | Which units, formats or scales changed?        | Before / after: _____     |
| Derivation  | Which values were calculated from others?      | Method: _____             |
| Exclusion   | Which records or conditions were left out?     | Reason: _____             |
| Final input | What exact dataset or values entered the tool? | Hash / version: _____     |

## Spot-check sample

| Sample | Original      | Working value   | Comparison        | Decision         |
|--------|---------------|-----------------|-------------------|------------------|
| 1      | Source: _____ | Prepared: _____ | Difference: _____ | Accept / correct |
| 2      | Source: _____ | Prepared: _____ | Difference: _____ | Accept / correct |
| 3      | Source: _____ | Prepared: _____ | Difference: _____ | Accept / correct |
| 4      | Source: _____ | Prepared: _____ | Difference: _____ | Accept / correct |

## Plan a Small Project Action Workbook

Use this page to turn a useful idea into owned steps and acceptance evidence. The aim is not to fill every blank; it is to expose the few facts that could change the decision. The evidence package should contain scope, owner, dependencies, estimate and done condition.

| Action field        | Prompt                                        | Record |
|---------------------|-----------------------------------------------|--------|
| Decision to support | What choice or action will this work inform?  | _____  |
| Audience            | Who must understand or approve it?            | _____  |
| Success measure     | What observable result would count as useful? | _____  |
| Decision date       | When is evidence needed, and why then?        | _____  |
| Critical assumption | Which belief could reverse the conclusion?    | _____  |
| Missing evidence    | What is the most important current gap?       | _____  |
| Stop condition      | What result requires pause or escalation?     | _____  |

## Outcome-specific plan

| Step         | Required action                                                  | Record                  |
|--------------|------------------------------------------------------------------|-------------------------|
| 1. Frame     | Restate plan a small project as a question                       | Owner / due: _____      |
| 2. Gather    | Collect scope, owner, dependencies, estimate and done condition  | Owner / due: _____      |
| 3. Calculate | Use inputs from user and task and data and consent               | Method / version: _____ |
| 4. Challenge | Test evaluation measures and failure and oversight               | Reviewer: _____         |
| 5. Decide    | Start only when the first milestone and stop condition are clear | Decision: _____         |

## One-sentence brief

We need to decide \_\_\_\_\_ by \_\_\_\_\_. The current evidence suggests \_\_\_\_\_, but the decision could change if \_\_\_\_\_. The next owner is \_\_\_\_\_.

## Assumption and Source Register

Source gate: an online result is not automatically authoritative. Record the original owner, document or dataset, publication or revision date, access date and the exact field used. Recheck time-sensitive values before publication or commitment.

| ID | Item                  | Evidence class               | Trace         | Decision impact  |
|----|-----------------------|------------------------------|---------------|------------------|
| A1 | User and task         | Fact / estimate / assumption | Source: _____ | Low / med / high |
| A2 | Data and consent      | Fact / estimate / assumption | Source: _____ | Low / med / high |
| A3 | Baseline method       | Fact / estimate / assumption | Source: _____ | Low / med / high |
| A4 | Model assumptions     | Fact / estimate / assumption | Source: _____ | Low / med / high |
| A5 | Evaluation measures   | Fact / estimate / assumption | Source: _____ | Low / med / high |
| A6 | Failure and oversight | Fact / estimate / assumption | Source: _____ | Low / med / high |

## Units, dates and versions

| Control           | Review question                               | Record            |
|-------------------|-----------------------------------------------|-------------------|
| Units             | Are symbols and conversion factors explicit?  | Checked by: _____ |
| Time basis        | Instant, daily, monthly, annual or lifecycle? | Basis: _____      |
| Currency / locale | Currency, tax location and rounding rule?     | Basis: _____      |
| Version           | Which tool, formula, policy or data revision? | Version: _____    |
| Applicability     | Why does the method fit this case?            | Reason: _____     |
| Exclusions        | What important factor is outside the model?   | Record: _____     |

## Assumption challenge

- What direct evidence would replace the largest assumption?
- Which assumption is both uncertain and capable of changing the decision?
- Would an independent person interpret every field the same way?
- Has a current authoritative source been checked for rates, limits, laws, standards or safety boundaries?
- Who is qualified to approve this use of the result?

## Scenario Comparison Matrix

Scenario work reveals whether plan a small project stays stable or depends on a narrow assumption. Change a driver deliberately, retain the same definitions and explain why each value is credible. Do not label an optimistic case as expected simply because it is preferred.

| Driver                | Low / adverse       | Working case    | High / favourable | Reason        |
|-----------------------|---------------------|-----------------|-------------------|---------------|
| User and task         | Conservative: _____ | Expected: _____ | Stretch: _____    | Source: _____ |
| Data and consent      | Conservative: _____ | Expected: _____ | Stretch: _____    | Source: _____ |
| Baseline method       | Conservative: _____ | Expected: _____ | Stretch: _____    | Source: _____ |
| Model assumptions     | Conservative: _____ | Expected: _____ | Stretch: _____    | Source: _____ |
| Evaluation measures   | Conservative: _____ | Expected: _____ | Stretch: _____    | Source: _____ |
| Failure and oversight | Conservative: _____ | Expected: _____ | Stretch: _____    | Source: _____ |

## Outcome comparison

| Comparison                    | Conservative        | Expected            | Stretch             |
|-------------------------------|---------------------|---------------------|---------------------|
| Calculated result             | _____               | _____               | _____               |
| Time implication              | _____               | _____               | _____               |
| Cost or resource implication  | _____               | _____               | _____               |
| Quality or performance effect | _____               | _____               | _____               |
| Risk and reversibility        | _____               | _____               | _____               |
| Evidence confidence           | Low / medium / high | Low / medium / high | Low / medium / high |

## Sensitivity note

The conclusion changes most when \_\_\_\_\_ moves from \_\_\_\_\_ to \_\_\_\_\_. That driver is supported by \_\_\_\_\_. Before commitment, we will \_\_\_\_\_.

## Verification and Decision Gate

**High-stakes gate:** Generated or predicted output can be wrong or biased. Protect private data and require human review for legal, medical, financial, safety or other consequential uses. For any consequential use, a qualified professional must verify the current authoritative source, inputs, method and limits. A completed workbook does not grant legal, financial, medical, engineering, laboratory, security, operational or safety approval.

| Gate               | Acceptance condition                                   | Status     | Evidence |
|--------------------|--------------------------------------------------------|------------|----------|
| Input trace        | Every critical value has a source, date and unit       | Pass / gap | _____    |
| Independent check  | A second method or reviewer could expose an error      | Pass / gap | _____    |
| Applicability      | The method fits the case and stated boundary           | Pass / gap | _____    |
| Range              | A plausible low and high case has been tested          | Pass / gap | _____    |
| Contrary evidence  | Conflicts are recorded and explained                   | Pass / gap | _____    |
| Privacy and rights | Sensitive and third-party material is controlled       | Pass / gap | _____    |
| Qualified review   | Required professional or authorised review is complete | Pass / gap | _____    |

## Decision record

| Decision field     | Choice                                  | Record           |
|--------------------|-----------------------------------------|------------------|
| Recommendation     | Proceed / revise / pause / reject       | Reason: _____    |
| Confidence         | Low / medium / high                     | Evidence: _____  |
| Conditions         | What must stay true?                    | Record: _____    |
| Reversal trigger   | What new evidence reopens the decision? | Record: _____    |
| Owner and approval | Who owns the decision and its limits?   | Name/date: _____ |

## Final rule

For plan a small project, start only when the first milestone and stop condition are clear. Record the decision separately from the calculation so a later reviewer can distinguish evidence, interpretation, approval and action.

## Handover, Reuse and Follow-Up

Close the record so another person can reproduce the work, understand the limits and continue plan a small project without relying on memory. Keep the smallest useful evidence package, protect sensitive information and give the final file a clear version.

| Handover field      | Record                             |
|---------------------|------------------------------------|
| Current conclusion  | _____                              |
| Evidence location   | Folder, record or reference: _____ |
| Open question       | _____                              |
| Decision condition  | _____                              |
| Next action / owner | _____                              |
| Review date         | _____                              |
| Retention / privacy | Keep until _____; access: _____    |

## Reuse checklist

| Reuse gate  | Question                                                          | Status       |
|-------------|-------------------------------------------------------------------|--------------|
| Definitions | Are the meanings of user and task and data and consent unchanged? | Yes / update |
| Sources     | Are rates, limits, data and guidance still current?               | Yes / update |
| Context     | Does the new case share the same scope and conditions?            | Yes / update |
| Method      | Is the chosen tool or method still applicable?                    | Yes / update |
| Review      | Has the right qualified or authorised person checked it?          | Yes / update |
| Version     | Has the reused record received a new date and identifier?         | Yes / update |

## 30-day learning loop

- Compare the expected result with what actually happened.
- Record whether an assumption, input definition or method caused avoidable rework.
- Improve one shared field or check; do not add process that supports no decision.
- Archive the final decision, not every unnecessary draft.

Original independent resource created for the ToolJolt companion library. This workbook is not copied from a manual, standard, website, training course or proprietary template, and it is not affiliated with or endorsed by any third-party product or authority. It is an educational planning and documentation aid, not professional advice, certification or permission to act. When adapting it, add only your own or properly licensed text and visuals; do not insert copied screenshots, logos, charts, artwork, source code, manuals or confidential material.