

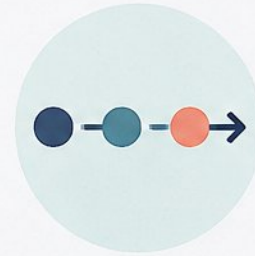
Documenting Decisions in Data Science

CLEARER THINKING. STRONGER OUTCOMES.



Benefits

- Reproducibility, auditability, knowledge retention, reduced rework



Lifecycle

- Lifecycle: framing, data selection, methodology, validation, deployment



Documentation layers

- Documentation layers: notes, decision records, model/data/system cards



Course path

- Course path: principles → templates → team workflows → hands-on practice



The Cost of Undocumented Decisions



Undocumented assumptions and silent cleaning choices hide real reasoning



Failed reproductions, compliance findings, lost revenue, and slow onboarding



Documentation debt compounds: stale or missing records erode trust in all docs



Training-serving feature divergence cost one team \$1.5M in lost revenue



AI-generated docs without human insight create volume without trust



Background and Regulatory Context

From informal notes to a more accountable and regulated future



- Shift from ad hoc notes to structured decision records in data teams
- EU AI Act Article 11 and Annex IV: technical documentation for high-risk systems, applying from August 2026
- High-risk rules phase in through 2027; medical and embedded AI rules follow into 2028
- SR 11-7, SS1/23, and FDA guidance set model documentation expectations in finance and health
- NIST AI RMF (Govern, Map, Measure, Manage) and ISO/IEC 42001 align documentation with responsible AI

Core Concepts and Vocabulary



Artifacts:

decision record, decision log, assumption register, model card, data card



Metadata:

context, options, rationale, evidence, owner, date, status, revisit condition



Atomic vs. composite decisions;
reversible vs. irreversible decisions

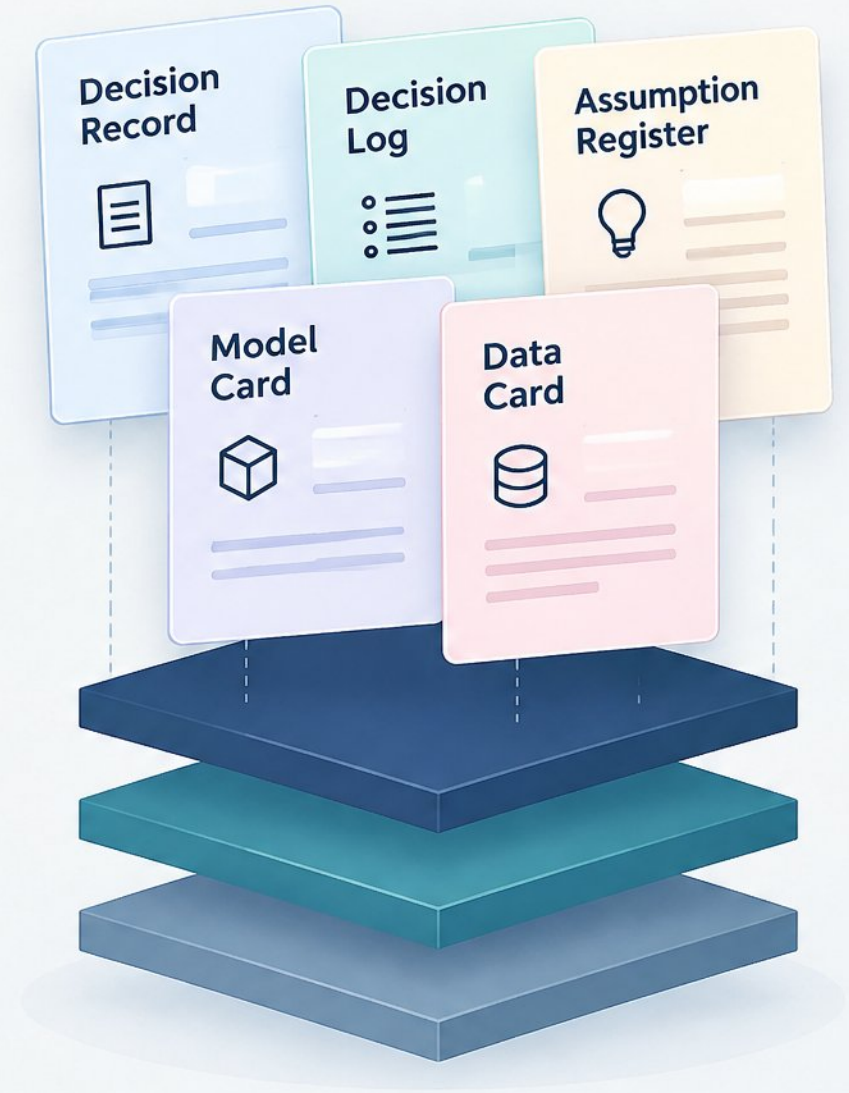


Traceability links decisions to data
versions, commits, experiments, model artifacts



Anti-patterns:

post hoc rationalization, decision theater,
documentation debt



When and What to Document

FOCUS ON WHAT MATTERS. BUILD A CLEARER PATH FORWARD.



Decision triggers: data inclusion, feature engineering, model selection, thresholds, deployment gating



Score decisions on impact, reversibility, uncertainty, and cross-team dependency



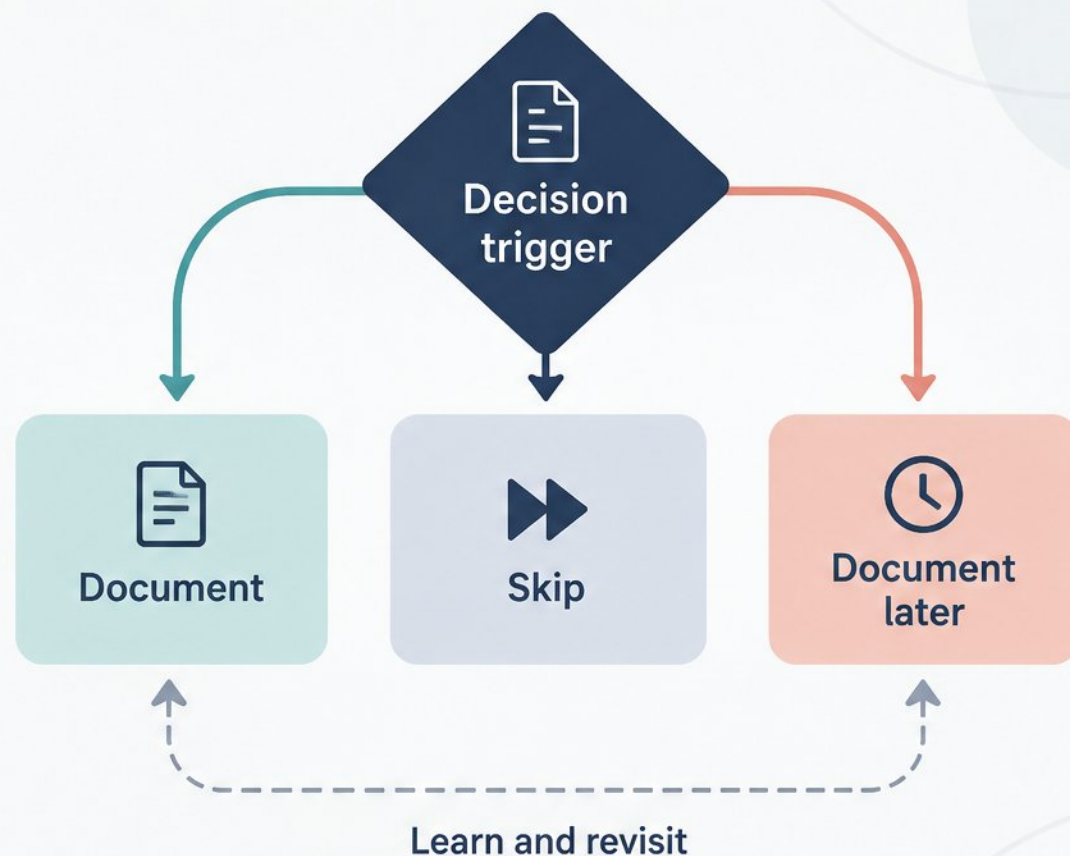
Document selectively: record baselines and direction-setting choices, skip routine steps



Capture negative results and abandoned paths to prevent duplicated effort



Assign clear roles: who writes, who reviews, who approves each record



**Clear roles
for each record**



Who writes



Who reviews



Who approves

Anatomy of a Decision Record



Standard fields: title, status, context, decision, consequences, alternatives, review trigger



Alternatives are the highest-value section — they stop re-litigating rejected options



Y-statement: "In context X, facing Y, I decided Z to achieve W, accepting V"



Record uncertainty: confidence levels, open questions, assumptions, revisit conditions



Link evidence: notebooks, dashboards, tickets, experiment IDs, data versions, commits

Decision Record

	Title
	Status
	Context
	Decision
	Consequences
	Alternatives
	Review trigger

Model Cards, Data Cards, and System Cards



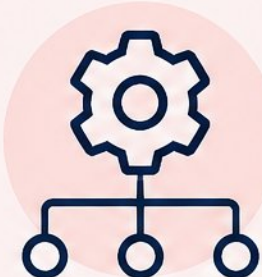
Model cards

Model cards: intended use, training data, metrics, limitations, safe-use guidance



Data cards and datasheets

Data cards and datasheets: provenance, collection methods, known biases, consent basis



System cards

System cards: public-facing summary of capabilities, limitations, design choices



These artifacts summarize what the system is; decision records explain why



Match artifact selection to team size, maturity, and regulatory exposure

Capturing Decisions During Analysis and Experimentation

FROM
CURIOSITY
TO REPRODUCIBLE
INSIGHTS

-  Document EDA choices: inclusion/exclusion, cleaning, transformations, imputation
-  Log hypotheses, results, learnings, and why the final approach won
-  Link decisions to experiment trackers (MLflow, W&B) and data catalogs
-  Analyst-inspector check: can a reviewer reproduce the workflow from docs alone?
-  LLM on the data path: record model version, prompts, sampling params, determinism



Practical Strategies for Teams



Embed decision records into sprint planning, design reviews, and model sign-off



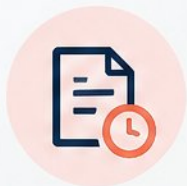
Create a shared decision log with consistent location and naming conventions



Automate capture: commits, run IDs, dataset versions, registry stage transitions



Scale from individual practice to team standard via champions and playbooks



For legacy projects: retroactive records, prioritized by risk



Tools and Workflow Integration

- Markdown decision records in Git: doc/adr/ or decisions/ directory
- Wiki templates (Confluence, Notion) for non-code decisions and stakeholders
- Model registry annotations and gated stage transitions enforce documentation
- Automated evidence: run IDs, dataset IDs, model hashes, approval events
- Choose tools by team size, maturity, regulatory needs—avoid tool sprawl



Review Triggers and Revisiting Decisions



Decisions decay: name the trigger that should force a revisit



Common triggers: cost threshold, vendor change, data availability



Supersede, never edit: link the old record to its replacement



Context drift: recheck original assumptions against current reality



Record outcomes: worked, mixed, or reversed—build a searchable history



Quality, Audit, and Governance



Quality criteria:

completeness, clarity, traceability, timeliness, honest trade-offs



Peer review and approval workflows:

independent review, effective challenge, clear sign-off



Audit readiness:

qualified third parties must replicate and understand your decisions



Governance models:

centralized vs. federated documentation ownership



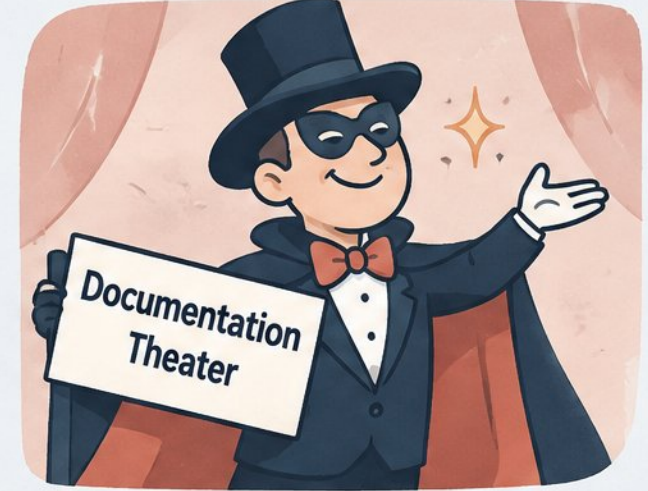
Track effectiveness:

onboarding time, re-litigation frequency, audit findings



Common Pitfalls and How to Avoid Them

RECOGNIZE | UNDERSTAND | TAKE ACTION



Documentation debt: missing, outdated, or inconsistent records compound silently



Post hoc rationalization: records written after the fact omit real trade-offs



Over-documentation and documentation theater: process slows delivery without adding value



Inconsistent formats: make search, review, and automation difficult



Cultural barriers: low incentives, time pressure, and fear of scrutiny



Counter with templates, incentives, and visible leadership support



Hands-On Workshop and Next Steps



Guided exercise: write a decision record for a recent analytical choice



Peer review using a quality checklist and analyst-inspector reproduction test



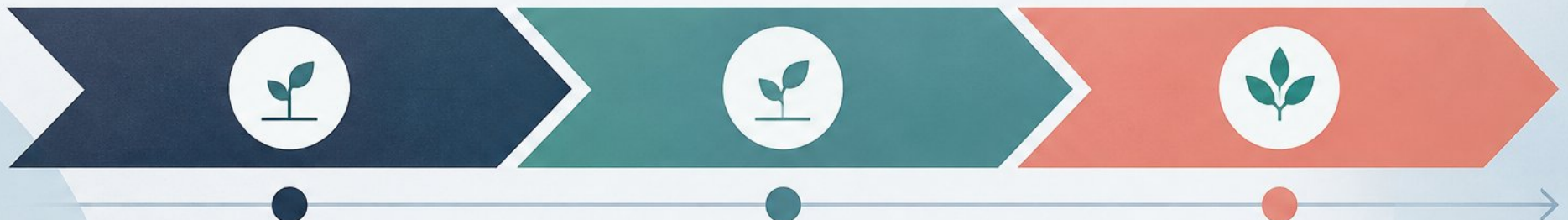
Adapt templates to your team's tools and regulatory environment



Build a 30-60-90 day adoption plan



Resources: templates, further reading, and internal champions



PersonWise.

PersonWise • AI digital-human course platform



Scan the code or click anywhere to watch this course live, with voice Q&A

personwise.ai/t/39cba025/public