

AGENT SPECIFICATION ENGINEERING



Building Reliable, Testable, and Governable
AI Systems with DSLs, Contracts, and
Executable Specifications

GEISON FLORES

Agent Specification Engineering

Building Reliable, Testable, and Governable AI Systems
with DSLs, Contracts, and Executable Specifications

Geison Goes

This book is available at <https://leanpub.com/agentspecificationengineering>

This version was published on 2026-07-28



This is a [Leanpub](#) book. Leanpub empowers authors and publishers with the Lean Publishing process. [Lean Publishing](#) is the act of publishing an in-progress ebook using lightweight tools and many iterations to get reader feedback, pivot until you have the right book and build traction once you do.

© 2026 Geison Goes

Contents

- Agent Specification Engineering 1**
 - Building Reliable, Testable, and Governable AI Systems with DSLs, Contracts, and Executable Specifications 1
- Copyright 2**
 - Code License 2
- Disclaimer 3**
- About the Author 4**
- Preface 5**
- How to Use This Book 6**
- Terminology and Notation 7**
- The Missing Engineering Layer Between Intent and Execution 8**
 - The Production Problem Is Larger Than Answer Quality 9
 - Prompting and Specification Solve Different Problems 10
 - Goals Are Not Instructions, and Autonomy Is Not Authority 11
 - A First AgentSpec 12
 - Executable Does Not Mean Fully Formal 13
 - The Reader Promise 14
- Part I – Why Agents Need Specifications 16**
- Chapter 1 – The Agent Reliability Problem 17**
 - The Demonstration Hides the State Space 17
 - Eight Reliability Failures Behind a Fluent Answer 18
 - Why a Better Prompt Cannot Close These Gaps 21
 - Unsafe Design: The Prompt as Control Plane 21
 - Corrected Design: Proposals Inside an Enforcement Shell 22

CONTENTS

Failure Semantics Must Be Part of the Product	23
Security and Privacy Are Execution Properties	24
Reliability and Operational Implications	24
When Not to Build an Agent	25
Design Review: A Fluent but Unreliable Support Agent	25
Production Checklist	27
Key Takeaways	27
Exercises	28
Chapter 2 – From Prompts to Behavioral Contracts	29
A Ladder of Explicitness	29
Design by Contract, Adapted Carefully	29
Preconditions Are Not Suggestions to Prepare Better	29
Postconditions Describe Observable Guarantees	29
Invariants Protect the Execution Across Steps	29
An Insufficient Structured Specification	29
A Contract-Oriented Refactoring	30
What Belongs in Prose, Data, Code, or Policy	30
Guardrails Are Not More Instructions	30
Security, Reliability, and Operations	30
Trade-offs and When Not to Formalize Further	30
Design Review: Refactoring an Operations Prompt	30
Production Checklist	30
Key Takeaways	31
Exercises	31
Chapter 3 – Anatomy of an Agent Specification	32
AgentSpec’s Design Boundary	32
The Minimal Specification	32
Metadata and Portability Make Context Reproducible	32
Vocabulary Prevents Domain Drift	32
Inputs and Outputs Define Ports, Not Merely Formats	32
Objectives Separate Direction from Termination	32
Capabilities and Tools Separate Ability from Adapters	33
The Contract Section Defines Cross-Cutting Obligations	33
State Turns a Loop into a Workflow	33
Policies Express Contextual Authority	33
Memory, Context, and Data Are Three Policies	33
Budgets and Resilience Define the Outer Failure Envelope	33

CONTENTS

Observability and Audit Serve Different Consumers	33
Evaluation Makes Release Expectations Executable	34
The Insufficient Specification Revisited	34
Implementation and Deployment Considerations	34
When AgentSpec Is Not the Right Representation	34
Design Review: Finding the Missing Section	34
Production Checklist	34
Key Takeaways	34
Exercises	35
Part II – Designing Agent Contracts	36
Chapter 4 – Goals, Outcomes, and Completion Conditions	37
A Vocabulary of Ending	37
Goals Need Bounds	37
Output Is Not Outcome	37
Acceptance Criteria Belong to Different Owners	37
Completion Must Be a Runtime Decision	37
Terminal States Are Semantically Different	37
Business Success Is Often Delayed	38
Model Confidence Is One Signal	38
An Unsafe Completion Design	38
Cancellation and Changing Intent	38
Security, Reliability, and Operational Implications	38
When Not to Use a Complex Outcome Model	38
Design Review: When “Resolved” Is Three Different Things	38
Production Checklist	39
Key Takeaways	39
Exercises	39
Chapter 5 – Tool Contracts and Side Effects	40
A Tool Is More Than a Function Signature	40
Typed Inputs and Outputs	40
Read, Reversible Write, Irreversible Write	40
Preconditions and Postconditions Around Effects	40
Timeouts Create Knowledge Problems	40
Designing Idempotency Keys	40
Retry Is a Classified Decision	41
Compensation Is Forward Recovery	41

CONTENTS

Human Approval Belongs Before the Effect	41
Tool-Result Provenance	41
An Unsafe and Corrected Tool Design	41
Security, Privacy, Reliability, and Operations	41
When Not to Expose a Tool	41
Design Review: The Timeout After Commit	42
Production Checklist	42
Key Takeaways	42
Exercises	42
Chapter 6 – State, Memory, and Context	43
Four Information Planes	43
Workflow State Is Not a Transcript	43
Domain State Has an Owner	43
Context Is a Purpose-Built Projection	43
Memory Needs a Retention Contract	43
Staleness Is a Semantic Property	43
Conflicting Context Must Remain Visible	44
Context Minimization and Privacy	44
An Unsafe and Corrected Design	44
Reliability and Operational Implications	44
When Not to Use Persistent Memory	44
Design Review: The Convenient Summary That Became Truth	44
Production Checklist	44
Key Takeaways	45
Exercises	45
Chapter 7 – Decisions and State Transitions	46
The Model Proposes; the State Machine Commits	46
States Should Represent Operational Meaning	46
Valid and Invalid Transitions	46
Guards Are Current Predicates	46
Transition Effects Need Their Own Safety	46
Terminal, Suspended, and Unknown States	46
Loops Need Progress Measures	47
Temporal Rules	47
Recovery and Versioning	47
An Unsafe and Corrected Workflow	47
Security and Operational Implications	47

CONTENTS

When Not to Use a Rich State Machine	47
Design Review: Two Workers, One Approval	47
Production Checklist	48
Key Takeaways	48
Exercises	48
Chapter 8 – Policies, Permissions, and Authority	49
Authority Is a Tuple, Not a Boolean	49
Least Privilege Starts With Architecture	49
Capabilities Are Delegable Units of Authority	49
Roles and Attributes Answer Different Questions	49
Deny by Default and Explicit Prohibition	49
Policy as Code	49
Tenant and Regional Boundaries	50
Approval Thresholds and Separation of Duties	50
Prompt Injection as an Authorization Problem	50
Confused Deputies	50
An Unsafe and Corrected Engineering Agent	50
Reliability and Operational Implications	50
When Not to Delegate Authority	50
Design Review: The Confused Deputy in a Repository Agent	51
Production Checklist	51
Key Takeaways	51
Exercises	51
Part III – Execution Architecture	52
Chapter 9 – Compiling Specifications into Workflows	53
Compilation as a Pipeline	53
Parsing and Normalization	53
Structural Validation	53
Semantic Validation	53
Binding to Deployment Reality	53
From State Model to Workflow Graph	53
Planning Is Constrained Graph Construction	54
Model-Facing Projections	54
Generated Runtime Artifacts	54
Runtime Execution	54
An Unsafe and Corrected Compiler	54

CONTENTS

Security and Supply-Chain Considerations	54
Operational Implications	54
When Not to Generate	55
Design Review: The Specification Passed but the Deployment Drifted	55
Production Checklist	55
Key Takeaways	55
Exercises	55
Chapter 10 – Deterministic and Probabilistic Boundaries	56
Five Responsibility Zones	56
What the Model Should Interpret	56
What Deterministic Code Should Validate	56
What Policy Engines Should Authorize	56
What Transaction Services Should Execute	56
What Humans Should Approve	56
The Deterministic Shell	57
Constrained Decoding and Schema Enforcement	57
Confidence Thresholds	57
False Determinism	57
An Unsafe and Corrected Decision Path	57
Latency and Cost Trade-offs	57
Security and Operational Implications	57
When Not to Use a Model	58
Design Review: JSON That Is Valid and Wrong	58
Production Checklist	58
Key Takeaways	58
Exercises	58
Chapter 11 – Orchestration, Delegation, and Multi-Agent Systems	59
Multi-Agent Is an Organizational Choice	59
Supervisor and Specialists	59
Delegation Contracts	59
Task Ownership	59
Authority Propagation	59
Shared State Without Shared Confusion	59
Handoff Schemas	60
Result Validation and Trust	60
Coordination Versus Decomposition	60
Circular Delegation and Termination	60

CONTENTS

Unsafe and Corrected Multi-Agent Design	60
Security and Privacy	60
Reliability and Operations	60
When Not to Use Multiple Agents	61
Design Review: When More Agents Create Less Accountability	61
Production Checklist	61
Key Takeaways	61
Exercises	61
Chapter 12 – Failure Semantics and Recovery	62
Failure Is a Contract Outcome	62
Transient and Permanent Failures	62
Invalid Model Output	62
Tool Timeout and Unknown Outcome	62
Partial Execution	62
Duplicate and Out-of-Order Results	62
Dependency Failure	63
Approval Timeout and Denial	63
Conflicting Decisions	63
Backoff and Jitter	63
Dead-Letter Handling	63
Compensation and Escalation	63
Retry Versus Safe Replay	63
Failure Budgets	64
Unsafe and Corrected Recovery	64
Security and Privacy	64
Operational Implications	64
When Not to Recover Automatically	64
Design Review: Recovery Without Wishful Rollback	64
Production Checklist	64
Key Takeaways	65
Exercises	65
Part IV – Verification and Quality	66
Chapter 13 – Testing Agent Specifications	67
A Layered Test Strategy	67
Schema Tests	67
Parser and Normalization Tests	67

CONTENTS

Semantic Tests	67
Contract Tests for Tools	67
State-Transition Tests	67
Invariant Tests	68
Property-Based Testing	68
Model-Based Testing	68
Simulation and Failure Injection	68
Adversarial Tests	68
Golden Datasets and Regression Suites	68
Mutation Testing	68
Test Doubles and Deterministic Models	69
Stochastic Test Strategy	69
Security and Operations	69
When Not to Use an End-to-End Model Test	69
Design Review: A Test Suite That Passed the Wrong System	69
Production Checklist	69
Key Takeaways	70
Exercises	70
Chapter 14 – Evaluating Behavior, Not Just Answers	71
Evaluation Begins With a Behavioral Record	71
Separate Hard Constraints From Quality	71
Response Quality	71
Task Completion	71
Tool Selection and Argument Correctness	71
Policy Compliance and Execution Order	71
Escalation Quality	72
Evidence Use and Source Attribution	72
Latency, Cost, and Reliability	72
Behavioral Scorecards	72
Offline Evaluation	72
Online Evaluation	72
Evaluation Leakage	72
Evaluator Disagreement	73
Unsafe and Corrected Evaluation	73
Security and Privacy	73
When Not to Use a Single Metric	73
Design Review: The 92 Percent Agent	73
Production Checklist	73

CONTENTS

Key Takeaways	73
Exercises	74
Chapter 15 – Guardrails and Runtime Enforcement	75
A Guardrail Must Control a Boundary	75
Input Validation	75
Output Validation	75
Domain Validation	75
Policy and Authorization	75
Rate Limits and Budgets	75
Sandboxing	76
Content Controls	76
Data-Loss Prevention	76
Approval Gates	76
Kill Switches and Capability Rollback	76
Layered Guardrails	76
Unsafe and Corrected Engineering Guardrails	76
Guardrail Composition and Failure	77
Security and Privacy	77
When Not to Add Another Guardrail	77
Design Review: The Guardrail That Lived in the Prompt	77
Production Checklist	77
Key Takeaways	77
Exercises	77
Chapter 16 – Observability, Auditability, and Replay	79
Three Evidence Products	79
Traces and Spans	79
Execution Timelines	79
Decision Records	79
Version Identity	79
Tool Calls and Policy Decisions	79
Provenance	80
Token and Cost Accounting	80
Audit Logs	80
Redaction	80
Correlation Identifiers	80
Safe Replay	80
Forensic Investigation	80

CONTENTS

Unsafe and Corrected Observability	81
Privacy Implications	81
When Not to Store a Transcript	81
Design Review: Reconstructing a Duplicate Effect	81
Production Checklist	81
Key Takeaways	81
Exercises	81
Part V – Production Patterns	83
Chapter 17 – Customer-Support Triage Agent	84
Scope and Architecture	84
AgentSpec Boundary	84
Data Model	84
State Machine	84
Tool Contracts	84
Policies and Privacy	84
Classification and Deterministic Rules	85
Testing Strategy	85
Walkthrough: The Suspicious-Login Ticket	85
Decision Table	85
Data-Threat Review	85
Observability and Audit	85
Failure Analysis	85
Unsafe Design	86
Corrected Design and Trade-offs	86
When Not to Use This Agent	86
Design Review: Mixed Intent Under Degraded Dependencies	86
Production Checklist	86
Key Takeaways	86
Exercises	87
Chapter 18 – Software-Engineering Agent	88
Scope and Architecture	88
AgentSpec Boundary	88
Repository Permissions	88
Sandboxing and Command Contracts	88
Planning Contract	88
Patch and Artifact Provenance	88

CONTENTS

Test Gates	89
Pull-Request Preparation and Human Review	89
Walkthrough: Protected Workflow Drift	89
State and Failure Paths	89
Evaluation and Tests	89
Observability	89
Trade-offs	89
When Not to Use This Agent	90
Design Review: A Correct Patch from an Unsafe Process	90
Production Checklist	90
Key Takeaways	90
Exercises	90
Chapter 19 – Financial-Operations Agent	91
Scope and Architecture	91
AgentSpec Boundary	91
Authoritative Data and Immutable Evidence	91
Threshold Policy	91
Separation of Duties	91
Review Package	91
State and Human Authorization	92
Idempotency and Duplicate Containment	92
Compensation and Reversal	92
Data Privacy and Explainability	92
Walkthrough: Threshold Drift	92
Tests and Evaluation	92
Observability and Audit	92
Failure Containment	93
Trade-offs and When Not to Use	93
Design Review: The Threshold at the Currency Cutoff	93
Production Checklist	93
Key Takeaways	93
Exercises	93
Chapter 20 – Multi-Agent Incident Response	94
Roles and Authority	94
Architecture	94
Supervisor AgentSpec	94
Immutable Incident Timeline	94

CONTENTS

Evidence Collection	94
Conflicting Hypotheses	94
Time Bounds and Budgets	95
Blast-Radius Control	95
Remediation State Machine	95
Communications Contract	95
Walkthrough: The Wrong-Region Hypothesis	95
Failure and Recovery	95
Testing and Evaluation	95
Observability and Post-Incident Replay	96
Trade-offs and When Not to Use	96
Design Review: Approval Invalidated by Late Telemetry	96
Production Checklist	96
Key Takeaways	96
Exercises	96
Chapter 21 – Portability, Evolution, and Governance	97
Model Portability	97
Framework Portability	97
Specification Versioning	97
Schema Evolution	97
Migration	97
Deprecation	97
Governance as an Engineering System	98
Review Processes	98
Agent Catalogs	98
Organizational Maturity	98
Production-Readiness Gates	98
Unsafe and Corrected Evolution	98
Security and Operational Implications	98
When Not to Pursue Portability	99
Design Review: A Portable Specification on an Incomplete Framework	99
Production Checklist	99
Key Takeaways	99
Exercises	99
From Agent Demos to Engineered Systems	100
The Discipline of Agent Specification Engineering	100
The Central Separations	100

CONTENTS

A Practical Adoption Sequence	100
A Ninety-Day Adoption Plan	101
What AgentSpec Contributes	103
The Standard for “Reliable”	103
The Engineering Model to Carry Forward	103
Appendix A – AgentSpec 0.1 Language Reference	104
Document and Serialization	104
Conformance Levels	104
Top-Level Inventory	104
metadata	104
agent	104
purpose and vocabulary	104
Data Ports: inputs and outputs	105
objectives	105
capabilities	105
tools	105
Predicates	105
contract	105
state	105
policies	106
memory	106
context	106
data	106
budgets	106
resilience	106
observability	106
audit	107
evaluation	107
portability	107
Validation Order	107
Minimum Runtime Events	107
Compatibility and Migration	107
Appendix B – AgentSpec JSON Schema	108
Dialect and Identity	108
Closed Top-Level Structure	108
Reusable Definitions	108
Required Versus Optional Fields	108

CONTENTS

Enumerations	108
Bounds	108
Arrays and Uniqueness	109
Embedded Port and Tool Schemas	109
Schema Validation in the Reference Engine	109
Versioning the Schema	109
Using the Schema in CI	109
Complete AgentSpec 0.1 Schema	109
Appendix C – Threat-Modeling Checklist	110
1. System and Purpose	110
2. Assets	110
3. Entry Points and Trust Boundaries	110
4. Identity and Delegation	110
5. Capability and Tool Authority	110
6. Prompt Injection and Untrusted Content	110
7. Data Privacy and Isolation	111
8. Context and Memory Integrity	111
9. Model and Provider Boundary	111
10. State and Workflow Integrity	111
11. Side Effects and Distributed Failure	111
12. Sandbox and Execution Environment	111
13. Human Approval and Social Engineering	111
14. Output and External Communication	112
15. Observability, Audit, and Replay	112
16. Supply Chain and Control Plane	112
17. Availability and Abuse	112
18. Evaluation and Detection	112
19. Residual Risk and Decision	112
Threat-Model Review Record	112
Appendix D – Production-Readiness Checklist	114
Release Header	114
Gate 1 – Purpose, Ownership, and Scope	114
Gate 2 – Specification and Build Integrity	114
Gate 3 – Identity, Authorization, and Secrets	114
Gate 4 – Data Governance and Privacy	114
Gate 5 – Tooling and Side-Effect Safety	114
Gate 6 – State, Concurrency, and Recovery	115

CONTENTS

Gate 7 – Model and Context Behavior	115
Gate 8 – Sandbox, Network, and Supply Chain	115
Gate 9 – Verification and Evaluation	115
Gate 10 – Observability and Audit	115
Gate 11 – Operations and Incident Response	115
Gate 12 – Rollout and Change Control	115
Release Decision	116
Fast Red-Flag Review	116
Appendix E – Evaluation Scorecards	117
Scorecard Header	117
Universal Hard Gates	117
Behavioral Rating Scale	117
Core Behavioral Dimensions	117
Weighting Template	118
Risk Slices	118
Customer-Support Triage Scorecard	119
Software-Engineering Agent Scorecard	119
Financial-Operations Agent Scorecard	119
Multi-Agent Incident Scorecard	119
Online Evaluation and Stop Conditions	119
Adjudication Protocol	119
Completed Scorecard Record	119
Appendix F – Exercise Solutions and Discussion Notes	121
Chapter 1 – The Agent Reliability Problem	121
Chapter 2 – From Prompts to Behavioral Contracts	121
Chapter 3 – Anatomy of an Agent Specification	121
Chapter 4 – Goals, Outcomes, and Completion Conditions	121
Chapter 5 – Tool Contracts and Side Effects	121
Chapter 6 – State, Memory, and Context	121
Chapter 7 – Decisions and State Transitions	122
Chapter 8 – Policies, Permissions, and Authority	122
Chapter 9 – Compiling Specifications into Workflows	122
Chapter 10 – Deterministic and Probabilistic Boundaries	122
Chapter 11 – Orchestration, Delegation, and Multi-Agent Systems	122
Chapter 12 – Failure Semantics and Recovery	122
Chapter 13 – Testing Agent Specifications	123
Chapter 14 – Evaluating Behavior, Not Just Answers	123

Chapter 15 – Guardrails and Runtime Enforcement	123
Chapter 16 – Observability, Auditability, and Replay	123
Chapter 17 – Customer-Support Triage Agent	123
Chapter 18 – Software-Engineering Agent	123
Chapter 19 – Financial-Operations Agent	123
Chapter 20 – Multi-Agent Incident Response	124
Chapter 21 – Portability, Evolution, and Governance	124
Appendix G – Glossary	125
Appendix H – Bibliography and Further Reading	126
Contracts, Formal Reasoning, and State	126
Data Formats, APIs, and Events	126
Security, Privacy, and Risk	126
Observability, Reliability, and Supply Chain	126
Language Models, Reasoning, and Tools	126
Suggested Reading Paths	126
Source Maintenance Note	127

Agent Specification Engineering

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Building Reliable, Testable, and Governable AI Systems with DSLs, Contracts, and Executable Specifications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Copyright

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Code License

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Disclaimer

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

About the Author

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Preface

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

How to Use This Book

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Terminology and Notation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Missing Engineering Layer Between Intent and Execution

The investigation agent appeared to work. Given a payment-dispute identifier, it retrieved a transaction, summarized the available evidence, and recommended that the merchant contest the claim. The answer was concise. The citations looked plausible. In a demonstration, the sequence felt almost indistinguishable from a well-designed product.

Then the production review began.

The transaction tool accepted a customer identifier, but no component checked that the identifier belonged to the tenant in the request. The model could request ninety days of history, or nine years, because the phrase “recent history” existed only in a prompt. A timeout from the evidence service caused the orchestration loop to repeat the entire plan. That was harmless while all tools were reads, but the next roadmap item would let the agent add a case note. Nobody had specified whether the note operation was idempotent. The final JSON matched its schema even when an evidence identifier came from a different dispute. The trace recorded model calls and tool payloads, yet it could not answer which policy version had authorized them. “Done” meant that the model stopped asking for tools.

None of these failures required an unintelligent model. They arose because the system had a mechanism for generating behavior but no authoritative description of the behavior permitted.

This is the gap between an agent demonstration and an engineered agent system. A prompt can communicate purpose, background, preferences, and examples. It is a valuable input to a probabilistic component. It is a poor substitute for an execution contract. The prompt does not own the database transaction, authenticate the caller, validate a state transition, reserve a budget, or prove that a timed-out side effect is safe to repeat. When teams make it responsible for those jobs, they move critical controls into a medium that the controlled component is free to reinterpret.

The discipline developed in this book starts with a different instruction to the engineering organization:

Do not prompt an agent to behave correctly. Specify the conditions under which correct behavior is possible.

That sentence does not dismiss prompts or promise perfect correctness. It identifies the proper control surface. A reliable system narrows the set of reachable behaviors before asking a model to choose among them. It represents authority outside the model, validates inputs and results at typed boundaries, constrains transitions through explicit guards, classifies failures before retrying, and records evidence that can be inspected without trusting the model's own account of what happened.

The Production Problem Is Larger Than Answer Quality

Much agent evaluation begins with the visible answer. Is it relevant? Is it factually accurate? Does it follow the requested format? These questions matter, but they cover only one layer of behavior.

Consider two executions that return the same accurate dispute recommendation. The first reads three authorized sources, cites their versions, stays within the tenant boundary, and stops when evidence conflicts. The second reads an unrelated customer's history, ignores a policy denial, calls a tool five extra times, and happens to produce the same sentence. An answer-only evaluator can mark both as correct. A production owner cannot.

Behavioral correctness includes the path, not only the terminal text. It asks whether the system selected an allowed tool, supplied valid and authorized arguments, observed the required order, respected time and cost limits, preserved invariants, used current evidence, escalated at the right boundary, and produced an auditable result. Sometimes the most correct behavior is to return no recommendation at all.

This distinction explains why “the model usually gets it right” is not a release criterion. A production system needs a defined population of cases, measurable properties, failure thresholds, and controls for the cases in which the model does not get it right. Stochastic quality and deterministic compliance can coexist, but they must not be confused.

The same reasoning applies to confidence. A model can be highly confident about a classification supported by stale or unauthorized data. It can be uncertain while the business rule is mechanically decisive. Confidence is evidence about a prediction process. It is not business certainty, permission, or proof.

Prompting and Specification Solve Different Problems

A prompt is an instruction-bearing message to a model. It can set a role, describe a goal, provide context, require a format, and show examples. Prompt engineering improves how a model interprets such a message. Specification engineering defines the system that surrounds the interpretation.

A specification answers questions that must remain true even if the model misunderstands the prompt:

- Which inputs are accepted, how are they classified, and which identity and tenant scopes must accompany them?
- Which capabilities exist, and which registered tools may exercise each capability?
- What preconditions must hold before a tool call, and what postconditions must hold after it?
- Which actions are explicitly prohibited?
- What state transitions are valid, and which guards or approvals authorize them?
- What ends the task? Which terminal states represent success, escalation, or failure?
- Which data may enter model context or persistent memory, and which source remains authoritative?
- Which retry classes are safe? Which mutating calls require idempotency keys? Which effects have explicit compensation?
- Which budgets limit iterations, tool calls, elapsed time, and cost?
- What must be traced, redacted, retained, and available for replay or audit?

Some of these answers will ultimately be expressed in code, policy bundles, database constraints, or infrastructure configuration. The specification does not replace those mechanisms. It provides a coherent, versioned source from which they can be generated, configured, checked, or reviewed.

The progression is therefore not “prompt, then more detailed prompt.” It is:

domain language $\square$ structured specification $\square$ executable behavioral contract $\square$ constrained agent execution $\square$ validation $\square$ observable and governable production system.

Each arrow represents an engineering transformation. Domain vocabulary removes ambiguity from concepts such as *dispute*, *evidence*, and *recommendation*. Structure makes fields inspectable. Contract semantics distinguish obligations from description. Runtime enforcement prevents invalid actions from becoming effects. Validation checks both shape and domain meaning. Observability and governance preserve the evidence required to operate and evolve the system.

Goals Are Not Instructions, and Autonomy Is Not Authority

Agent discussions often blur a goal with the actions used to pursue it. “Resolve the customer’s issue” is an outcome-oriented goal. “Search the account, edit the subscription, apply a credit, and close the ticket” is a sequence of instructions. The goal can remain stable while the valid instructions change with policy, product state, region, customer segment, or incident conditions.

An autonomous component may choose among allowed actions without asking a human to select every step. That autonomy does not imply unlimited authority. A build agent can decide which tests to run while lacking permission to merge its own pull request. An incident-analysis agent can select telemetry queries while lacking permission to restart a service. A financial-review agent can recommend an operation while being structurally incapable of executing it.

This separation is a feature, not a limitation. It lets teams increase autonomy within a bounded capability set without expanding the blast radius of a model error. It also makes human approval meaningful. Approval becomes a scoped state transition backed by evidence, an approver role, expiration, and an explicit effect. It is not a vague request for someone to “check the AI.”

Tool availability must be separated in the same way. Registering a tool tells the runtime how an adapter can be invoked. Permission answers whether this subject may invoke that action on this resource for this purpose under

current conditions. If a prompt-injected document says “issue a refund,” the authorization decision should remain unchanged. Treating prompt injection only as a contest between trusted and untrusted words leaves the model responsible for privilege enforcement. Treating it as an authorization problem ensures that malicious text cannot manufacture capability.

A First AgentSpec

AgentSpec 0.1 is the book’s pedagogical, semi-formal domain-specific language for expressing these distinctions. It is not proposed as an industry standard. It exists so that the concepts can remain consistent across prose, code, diagrams, tests, and case studies.

The following deliberately small fragment describes the outer boundary of the Payment Dispute Investigation Agent:

```
1  metadata:
2    id: payments.dispute_investigator
3    version: 1.0.0
4    agentspec_version: "0.1"
5
6  agent:
7    identity: payments.dispute_investigator
8    risk_class: high
9
10 capabilities:
11   - retrieve_transaction
12   - retrieve_dispute_evidence
13   - recommend_action
14
15 contract:
16   prohibited_actions:
17     - move_money
18     - issue_refund
19     - modify_customer_data
20     - close_dispute
21   output_validation: strict
22
23 policies:
24   default: deny
```

Even this fragment makes claims that a prompt commonly leaves implicit. The agent has a stable identity and a versioned contract. Its risk class is high. Its

capabilities are enumerated. Several tempting actions are outside its authority. Unknown access defaults to denial.

The fragment is not sufficient for production. It lacks typed inputs and outputs, tool contracts, tenant conditions, state transitions, evidence rules, approvals, budgets, retry behavior, and audit requirements. That insufficiency is useful because it gives reviewers a concrete artifact to reject. “The prompt seems comprehensive” is difficult to falsify. “No completion condition is declared” is a precise defect.

The complete specification in the companion repository adds those missing sections. Its tools are read-only. Each tool has argument and result schemas, a timeout, an authorization rule, and predicates evaluated before and after invocation. The agent may produce a non-binding recommendation, but a reviewer approval gate controls publication. Tenant isolation and the prohibition on financial mutation are runtime invariants. The execution has bounded model calls, tool calls, iterations, elapsed time, and cost units. Audit events are correlated while sensitive fields are redacted.

The reference engine loads this document, validates its structure against JSON Schema Draft 2020-12, performs cross-field semantic checks, and enforces it around a deterministic test model. The implementation is intentionally compact, but its tests demonstrate an essential property: a model decision is a proposal until deterministic code validates and authorizes it.

Executable Does Not Mean Fully Formal

The term *executable specification* can suggest mathematical proof or automatic code generation from a complete formal model. Those techniques are valuable in the right systems, but they are not prerequisites for this book.

AgentSpec is semi-formal. Its document shape is machine-validated. Many predicates, state transitions, budgets, permissions, and output schemas are executable. Some prose fields retain human meaning. Some properties require domain services or policy engines. Some important concerns, including whether a purpose is ethically justified, cannot be settled by a schema.

This mixed representation is deliberate. Mature engineering places a rule in the medium capable of enforcing or communicating it:

- prose for rationale, trade-offs, and meaning that requires judgment;

- structured data for versioned declarations and portable review;
- deterministic code for validation, resource accounting, and state enforcement;
- policy for contextual authorization and organizational rules;
- model inference for interpretation, classification, synthesis, and choices that cannot be reduced to stable algorithms;
- human approval for consequential authority and accountable judgment.

False determinism is as dangerous as missing determinism. Converting an uncertain model score to a boolean threshold does not make the underlying claim certain. A schema-valid rationale can still cite irrelevant evidence. A state machine can enforce a flawed business process perfectly. Specification engineering improves the system's inspectability and constraint; it does not eliminate the need for evidence, evaluation, and review.

The Reader Promise

By the end of this book, you will be able to review an agent as a production system rather than as an impressive conversation. You will be able to separate model responsibilities from runtime, policy, transactional, and human responsibilities. You will know how to write tool contracts that account for effects, timeouts, idempotency, retries, and compensation. You will model state without confusing it with memory or context. You will design delegation so that authority narrows or remains explicit across handoffs. You will test behavioral properties and evaluate execution paths, not only final answers. You will build traces and audit evidence that support investigation and safe replay without turning logs into a sensitive-data archive.

The four complete case studies are designed to prevent the method from remaining abstract. Customer-support triage emphasizes privacy, routing, and escalation. Software-engineering work emphasizes sandboxing, repository permissions, test gates, artifact provenance, and human review. Financial operations emphasizes thresholds, immutable evidence, separation of duties, and failure containment. Multi-agent incident response emphasizes coordination, conflicting hypotheses, time bounds, commander authority, and blast-radius control.

The practical adoption sequence is incremental. A team can begin by inventorying capabilities and prohibited actions, then add typed tool contracts,

explicit completion, state transitions, policy checks, failure classifications, behavioral tests, and audit evidence. It does not need to replace its model or framework to start. The specification layer is valuable precisely because it sits between domain intent and execution technology.

The first task is to understand the failure surface that this layer must control.

Part I – Why Agents Need Specifications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 1 – The Agent Reliability Problem

At 09:14, the support agent received a message from a customer who could not access a paid feature. At 09:15, it fetched the account, found a recently expired entitlement, and proposed a one-month courtesy extension. At 09:16, it called the entitlement tool and replied that access had been restored.

The transcript looked exemplary. The production record did not.

The customer belonged to a reseller-managed tenant whose contract prohibited direct entitlement changes. The account response included that fact, but it appeared in an unstructured note among dozens of fields. The tool adapter authenticated the service account, not the end user's delegated scope. The extension call timed out after the remote service committed the change but before it returned a response. The agent retried, creating a second extension. Its final message reported one month; the authoritative system now recorded two. A nightly reconciliation job caught the discrepancy eleven hours later.

This is a composite scenario, not a report about a particular company. Every component could have behaved according to its local design. The model chose a plausible action. The tool accepted a structurally valid request. The remote service performed the requested mutation. The retry mechanism reacted to a timeout. The answer described the intended result. The system still violated policy and produced an incorrect effect.

The chapter's engineering question is: **which failure classes appear when a probabilistic decision-maker is connected to stateful tools, and why are they not solved by better instructions alone?**

The Demonstration Hides the State Space

A demonstration usually follows a cooperative path. The input is well formed. Dependencies respond promptly. The model selects the expected tool. The

tool returns the expected shape. The task reaches an obvious end. Observers see the system’s capability along one trajectory.

Production exposes a state space. Inputs are incomplete, adversarial, duplicated, stale, or valid under a different tenant. Dependencies fail before, during, or after applying an effect. Results arrive out of order. Policies change while work is suspended. A user cancels after an action begins. A human approval expires. The model repeats a plan, invents an argument, or interprets the same context differently after an upgrade. Two executions race on the same domain record.

The important difference is not simply scale. It is the number of transitions whose semantics must be defined. If the design specifies only the successful path, every unmodeled branch inherits accidental behavior from the model, orchestration library, adapter, or remote service.

Traditional software has the same underlying problem, but agents make the omission easier to miss. A procedural implementation reveals many branches as code. An agent loop can compress them into a deceptively short pattern:

```
1 while not done:
2     decision = model.decide(context, tools)
3     result = call(decision.tool, decision.arguments)
4     context.append(result)
```

This fragment does not state what done means. It does not distinguish an available tool from an authorized tool. It does not classify effects, validate domain relationships, limit iterations, pin authoritative state, or define what a timeout means. The absence of code is not the absence of behavior. It is behavior delegated to defaults.

Eight Reliability Failures Behind a Fluent Answer

The following failure classes recur across agent domains. They overlap, and a single incident commonly crosses several.

Ambiguous goals

“Resolve the issue” can mean restore service, answer a question, create a case, offer a workaround, or change an entitlement. Optimizing for resolution

without defining allowed outcomes encourages the system to treat every reachable action as a means to the goal. A goal supplies direction; it does not supply authority or completion semantics.

Ambiguity also appears in apparently measurable goals. “Minimize handling time” does not say whether escalation counts as failure, whether waiting for approval stops the clock, or whether closing an unresolved case improves the metric. An agent can optimize the literal measure while violating the intended business outcome.

Uncontrolled tool use

A tool catalog describes capability. When the catalog is passed directly to a model, the model often becomes the de facto authorization layer. It decides whether a tool is relevant, but relevance and permission are different questions.

The problem is not limited to obviously dangerous tools. A read tool can expose another tenant’s data, expand context beyond necessity, or retrieve a restricted field. A search tool can become an exfiltration channel. A harmless-looking “add note” operation can trigger downstream automation. Effect classification must include domain consequences, not only whether an HTTP method is GET or POST.

Hidden and conflicting state

Agent frameworks commonly maintain a conversation, scratchpad, or memory object. Domain services maintain accounts, tickets, payments, repositories, and incidents. Workflow engines maintain execution progress. These are different kinds of state with different owners and consistency rules.

If a model’s context says a case is open while the case service says it was closed, the system needs a conflict policy. If memory contains an old customer preference, the runtime needs a freshness and provenance rule. Treating all retained text as “what the agent knows” hides which record is authoritative.

Malformed or semantically invalid arguments

Structured output reduces syntax errors, but schema validity is only the first boundary. An amount can be a valid number and exceed the caller’s limit. A

repository path can be a valid string and escape the authorized workspace. A dispute identifier and transaction identifier can each be valid while referring to different cases.

These are semantic constraints. They require cross-field predicates, authoritative lookups, policy decisions, or transactional checks. When a system celebrates that the model “always returns valid JSON,” it may be measuring the least consequential property.

Silent partial failure

An execution can retrieve two of three sources, summarize the partial set, and present the result as complete. A multi-step mutation can apply the first effect and fail before the second. A communications agent can publish an internal update while the public update fails. Without explicit completion conditions and failure states, partial success resembles success in the transcript.

Silence is especially dangerous when adapters normalize errors into prose. If a tool returns `{"message": "service unavailable"}` under a permissive output schema, the model may interpret the message as data rather than a failed dependency. Failure must have a typed representation outside natural-language interpretation.

Unbounded iteration

An agent that can plan, inspect results, and replan has a loop. Every loop needs a variant or a bound: a reason it progresses, a maximum number of iterations, or both. “Stop when the task is complete” is circular when completion is itself a model judgment.

Unbounded iteration consumes latency and money, but the operational risk is broader. Repeated reads can hit rate limits. Repeated writes can duplicate effects. Replanning can drift from the initial scope. Context can grow until early constraints receive less effective attention. Budgets should cover model calls, tool calls, elapsed time, cost, and domain-specific resources.

Authority without policy

A service credential often has more privilege than the user or task requires. If the agent inherits the credential's reach, it becomes a confused deputy:

an authorized component induced to use its authority for an unauthorized purpose. Text from a retrieved document, tool result, issue, or web page can influence model decisions. It must not change the authorization context.

Least privilege is not achieved by telling the model to be cautious. It requires credentials, scopes, resource checks, sandbox boundaries, and policy decisions that remain effective even when the model proposes a prohibited action (Saltzer and Schroeder, 1975).

Non-reproducible behavior

The same user input is not the same execution when the model version, system prompt, retrieved context, tool data, policy bundle, specification, or adapter changes. Temperature is only one source of variation. Even a deterministic model setting cannot recreate a dependency’s past state.

Reproducibility therefore means more than resubmitting a transcript. A useful execution record identifies versions, records decisions and relevant results, preserves correlation, and states which side effects may be simulated or replayed. Replay is an explicit safety mode, not a synonym for retry.

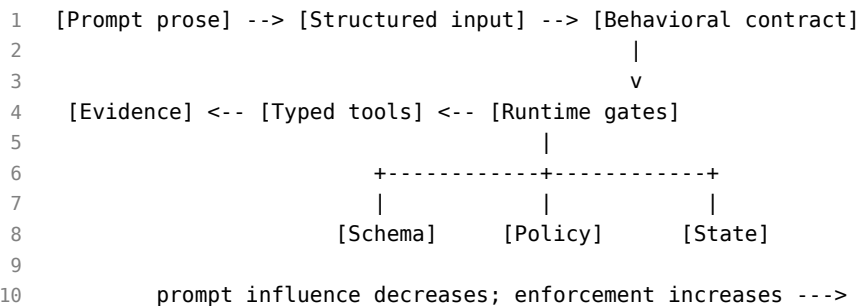


Figure 1. A specification moves consequence-bearing decisions out of prompt prose and into independent enforcement layers.

Why a Better Prompt Cannot Close These Gaps

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Unsafe Design: The Prompt as Control Plane

The following design is common because it is easy to assemble:

1. Build a broad system prompt containing operational rules.
2. Register every tool the product may eventually need.
3. Give the model the current conversation and selected data.
4. Execute any tool call whose arguments match the tool's JSON schema.
5. Append tool results to context.
6. Repeat until the model returns a final answer.

Its central flaw is not that the prompt is poorly written. The flaw is that model output crosses directly into effects. The model performs interpretation, planning, authorization, transition selection, error classification, and completion judgment. Each responsibility is probabilistic, and the runtime provides few independent grounds for refusal.

Adding a second model as a critic does not change the control architecture. The critic may improve detection, but it remains another probabilistic proposal unless a deterministic or accountable authority acts on its output under explicit rules. A “safety agent” with access to the same ambiguous context can disagree without resolving which component is authoritative.

Corrected Design: Proposals Inside an Enforcement Shell

The corrected architecture treats the model as a proposer within a deterministic shell.

The runtime first validates the request and constructs an authorization context from authenticated identity, tenant, purpose, resource scope, and roles. It loads a reviewed specification version and verifies that registered adapters match declared tool contracts. Only then does it ask the model for a decision.

A proposed tool call crosses several independent checks:

1. The tool is declared by the specification.
2. The agent has the corresponding capability.

3. No explicit deny policy matches the action.
4. An allow rule authorizes this subject, resource, purpose, and context.
5. Arguments match the input schema and cross-field domain rules.
6. Preconditions hold against the current workflow and authoritative domain state.
7. Budgets permit another call.
8. Required approval and idempotency evidence is present.

After the call, the runtime classifies transport and domain outcomes, validates the result schema, checks postconditions and invariants, records an audit event, and permits only a declared next transition. The model sees the normalized result, not necessarily the raw sensitive payload.

This shell cannot guarantee that a recommendation is wise. It can guarantee that an undeclared refund tool is not executed by this runtime, that a tenant mismatch fails closed, that a terminal state is reached only through declared transitions, and that an output lacking required evidence is rejected. Those properties are narrower than intelligence and far more operationally useful.

Failure Semantics Must Be Part of the Product

Failures are not exceptions to the agent design. They are states in it.

A timeout is not a single semantic category. A read timeout can often be retried under a bounded policy. A mutation timeout may leave the outcome unknown: the remote service could have committed before the response was lost. Retrying the mutation is safe only when the operation is idempotent under the same key or when the runtime can prove the first attempt was not applied. HTTP explicitly cautions against automatically retrying non-idempotent requests without such knowledge (IETF, 2022).

Likewise, compensation is not rollback. A compensating action occurs later in a world that may have changed. Reversing an entitlement extension can trigger a notification, interact with a customer action, or violate a new policy. The Saga pattern formalized compensation for long-lived transactions, but it never made external reality transactional (Garcia-Molina and Salem, 1987). Agent designs should name compensating tools and their preconditions instead of assuming every write has an inverse.

An actionable failure model distinguishes at least:

- transient dependency failure, eligible for bounded retry;
- permanent dependency failure, eligible for escalation or terminal failure;
- invalid model output, eligible for a limited repair attempt or failure;
- policy denial, not eligible for prompt-based persuasion;
- precondition conflict, eligible for state refresh and replanning within scope;
- unknown mutation outcome, eligible for status reconciliation, not blind retry;
- partial execution, eligible for explicit compensation or human recovery;
- budget exhaustion, eligible for a bounded summary and escalation;
- approval timeout or denial, eligible for a declared suspended or terminal path.

This vocabulary makes operational metrics meaningful. “Agent error rate” is too coarse if it combines policy working as designed, unavailable dependencies, invalid model outputs, and duplicated effects.

Security and Privacy Are Execution Properties

Prompt injection demonstrates why content and authority must be separated. An issue description, email, document, or web page can contain imperative text. The model may need to interpret that text as data. It must not inherit the author’s privileges.

The runtime should construct authority from authenticated control-plane inputs, not from model context. Retrieved content cannot add roles, expand resource scopes, register tools, alter prohibited actions, or approve transitions. A model proposal that names a forbidden action should produce a policy denial and security signal, not a debate inside the prompt.

Privacy uses the same boundary. Context minimization is not merely a token optimization. The system should retrieve only fields needed for the declared purpose, enforce tenant and regional handling rules, prevent restricted values from entering model input when required, redact telemetry, and expire retained memory according to policy. An audit log that stores every raw prompt and tool payload may increase forensic convenience while creating a more dangerous secondary dataset.

Reliability and Operational Implications

Explicit contracts change how a team operates the system.

Release review can diff a permission change independently from prompt prose. On-call engineers can see whether latency came from the model, a policy decision, an approval wait, or a tool. A policy denial can be counted separately from a dependency failure. A specification version can be rolled back without changing the model. A replay environment can substitute recorded read results while preventing mutations. Product owners can review completion conditions in domain language.

The discipline introduces costs. Specifications must be maintained. Adapters need accurate effect and schema declarations. Policy and domain state need clear ownership. Strong validation can reject cases that an unconstrained model might have handled acceptably. State migrations complicate long-running executions. Audit evidence requires retention and access controls.

These costs should be compared with the alternative: rules distributed across prompts, adapter code, orchestration defaults, tribal knowledge, and incident runbooks. Specification engineering does not create the obligations. It makes them visible.

When Not to Build an Agent

An agent is unnecessary when a stable deterministic procedure solves the problem more cheaply and safely. If inputs are structured, decision rules are complete, and the valid workflow has little ambiguity, ordinary application code or a workflow engine is often the better design.

An agent is also a poor choice when the organization cannot define ownership, authoritative state, permissions, or failure handling for the underlying process. Adding a model will not repair a process that nobody can specify. A read-only assistant with no tool authority may still be useful, but it should be described honestly as a different product.

Finally, do not grant production effects merely to demonstrate autonomy. Start with analysis and recommendations, collect behavioral evidence, and add narrowly scoped effects only when the organization can enforce and operate them.

Design Review: A Fluent but Unreliable Support Agent

Consider a support agent that reads a conversation, looks up the account, adds an internal note, and returns a suggested reply. Its pilot shows high reviewer preference, yet production review finds three unexplained duplicate notes and one reply built from another customer's order. The first design question is not whether a larger model improves accuracy. It is which execution decisions were never made explicit.

Draw the path from authenticated request to output. The model receives a customer ID copied from message metadata, chooses a lookup tool, and supplies an order ID extracted from prose. The tool gateway validates both as strings, but neither gateway nor order service verifies that the order belongs to the authenticated tenant and customer. The model then decides that note creation succeeded because the adapter timed out after it sent the request. A generic retry wrapper calls again with a new request ID. Finally, the loop stops when the model emits a response object with `done: true`.

The corrected boundary begins by deriving tenant and customer scope from authenticated identity, not from model text. The account/order service checks the relationship again. `note.create` receives a stable operation key bound to case, message version, note digest, and purpose. Timeout produces `outcome_unknown`; status reconciliation decides whether the note exists before any retry. The workflow runtime, rather than the model, commits `draft_ready` only after required reads are current, the note receipt is authoritative, and the draft schema/evidence checks pass. `escalated` and `failed` remain legitimate terminal outcomes.

Now revisit the attractive pilot metric. Reviewer preference measured the prose after a mostly successful path. It did not exercise guessed identifiers, stale sessions, duplicate delivery, provider timeout after commit, or policy denial. Add those cases, but do not rely on evaluation to enforce the boundary. Even if an adversarial case convinces the model to use a different customer ID or declare success, service authorization and workflow predicates must reject it. The evaluation measures how often such proposals occur and whether escalation is useful; runtime controls determine what becomes real.

The evidence package for this review should include a data-flow diagram, specification and tool registry, identity mapping, timeout trace, duplicate-note reproduction, isolation test, and one forensic reconstruction. If the team

cannot reconstruct which proposal, policy decision, operation key, service receipt, and state transition produced a note, the system is not yet operationally specified. The lesson is practical: reliability improves when every consequence-bearing decision has a named owner and an observable failure mode.

Production Checklist

- Is the agent's purpose distinct from its allowed actions?
- Are success criteria and completion conditions machine-checkable where practical?
- Is every tool classified by read, reversible write, or irreversible write effect?
- Does authorization bind subject, tenant, purpose, action, and resource?
- Are tool arguments validated beyond JSON shape?
- Are workflow state, model context, memory, and authoritative domain state separated?
- Are valid, suspended, failed, and terminal transitions declared?
- Are model calls, tool calls, iterations, elapsed time, and cost bounded?
- Are retryable failures distinguished from unknown mutation outcomes?
- Do mutating retries require idempotency or proof of non-application?
- Are compensating actions explicit and treated as forward recovery?
- Can retrieved content influence a decision without expanding authority?
- Are specification, policy, model, prompt, tool, and domain-record versions observable?
- Can an operator reconstruct the execution without exposing secrets in logs?
- Is there a declared response for policy denial, approval timeout, and partial execution?

Key Takeaways

An agent demo shows one trajectory; production requires semantics for a state space. Fluent output does not prove correct tool use, authority, execution order, or effects. Structured output solves syntax before it solves domain correctness. Autonomy should select among permitted actions, not define

permission. Reliability comes from deterministic enforcement around probabilistic proposals, with explicit failure and evidence models.

The next chapter develops the conceptual bridge from prompts and schemas to executable behavioral contracts.

Exercises

1. Take an agent loop from a system you know. List every decision currently delegated to the model or framework default. Classify each as interpretation, planning, validation, authorization, effect execution, state transition, failure handling, or completion.
2. Write three different completion conditions for “resolve a support issue.” Identify which one can be gamed by premature closure and how you would detect it.
3. For a tool that creates a calendar event, classify the outcomes of a connection timeout. Design an idempotency key and a status-reconciliation step.
4. Analyze a read-only repository agent. Identify at least four security or privacy failures that remain possible without mutating tools.
5. Rewrite the unsafe entitlement prompt as a division of responsibilities among prompt, schema, policy, domain service, state machine, and human approval.
6. Design a failure taxonomy for an agent that prepares, but cannot send, customer communications. Explain which failures should be retried, escalated, or treated as expected policy behavior.

Chapter 2 — From Prompts to Behavioral Contracts

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

A Ladder of Explicitness

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design by Contract, Adapted Carefully

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Preconditions Are Not Suggestions to Prepare Better

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Postconditions Describe Observable Guarantees

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Invariants Protect the Execution Across Steps

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

An Insufficient Structured Specification

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

A Contract-Oriented Refactoring

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

What Belongs in Prose, Data, Code, or Policy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Guardrails Are Not More Instructions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security, Reliability, and Operations

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Trade-offs and When Not to Formalize Further

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: Refactoring an Operations Prompt

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 3 — Anatomy of an Agent Specification

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

AgentSpec's Design Boundary

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Minimal Specification

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Metadata and Portability Make Context Reproducible

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Vocabulary Prevents Domain Drift

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Inputs and Outputs Define Ports, Not Merely Formats

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Objectives Separate Direction from Termination

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Capabilities and Tools Separate Ability from Adapters

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Contract Section Defines Cross-Cutting Obligations

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

State Turns a Loop into a Workflow

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Policies Express Contextual Authority

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Memory, Context, and Data Are Three Policies

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Budgets and Resilience Define the Outer Failure Envelope

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Observability and Audit Serve Different Consumers

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Evaluation Makes Release Expectations Executable

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Insufficient Specification Revisited

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Implementation and Deployment Considerations

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When AgentSpec Is Not the Right Representation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: Finding the Missing Section

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Part II – Designing Agent Contracts

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 4 — Goals, Outcomes, and Completion Conditions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

A Vocabulary of Ending

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Goals Need Bounds

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Output Is Not Outcome

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Acceptance Criteria Belong to Different Owners

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Completion Must Be a Runtime Decision

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Terminal States Are Semantically Different

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Business Success Is Often Delayed

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Model Confidence Is One Signal

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

An Unsafe Completion Design

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Cancellation and Changing Intent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security, Reliability, and Operational Implications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Use a Complex Outcome Model

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: When “Resolved” Is Three Different Things

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 5 — Tool Contracts and Side Effects

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

A Tool Is More Than a Function Signature

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Typed Inputs and Outputs

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Read, Reversible Write, Irreversible Write

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Preconditions and Postconditions Around Effects

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Timeouts Create Knowledge Problems

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Designing Idempotency Keys

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Retry Is a Classified Decision

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Compensation Is Forward Recovery

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Human Approval Belongs Before the Effect

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Tool-Result Provenance

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

An Unsafe and Corrected Tool Design

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security, Privacy, Reliability, and Operations

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Expose a Tool

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: The Timeout After Commit

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 6 — State, Memory, and Context

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Four Information Planes

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Workflow State Is Not a Transcript

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Domain State Has an Owner

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Context Is a Purpose-Built Projection

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Memory Needs a Retention Contract

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Staleness Is a Semantic Property

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Conflicting Context Must Remain Visible

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Context Minimization and Privacy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

An Unsafe and Corrected Design

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Reliability and Operational Implications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Use Persistent Memory

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: The Convenient Summary That Became Truth

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 7 — Decisions and State Transitions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Model Proposes; the State Machine Commits

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

States Should Represent Operational Meaning

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Valid and Invalid Transitions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Guards Are Current Predicates

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Transition Effects Need Their Own Safety

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Terminal, Suspended, and Unknown States

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Loops Need Progress Measures

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Temporal Rules

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Recovery and Versioning

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

An Unsafe and Corrected Workflow

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security and Operational Implications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Use a Rich State Machine

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: Two Workers, One Approval

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 8 — Policies, Permissions, and Authority

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Authority Is a Tuple, Not a Boolean

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Least Privilege Starts With Architecture

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Capabilities Are Delegable Units of Authority

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Roles and Attributes Answer Different Questions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Deny by Default and Explicit Prohibition

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Policy as Code

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Tenant and Regional Boundaries

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Approval Thresholds and Separation of Duties

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Prompt Injection as an Authorization Problem

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Confused Deputies

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

An Unsafe and Corrected Engineering Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Reliability and Operational Implications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Delegate Authority

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: The Confused Deputy in a Repository Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Part III — Execution Architecture

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 9 — Compiling Specifications into Workflows

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Compilation as a Pipeline

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Parsing and Normalization

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Structural Validation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Semantic Validation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Binding to Deployment Reality

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

From State Model to Workflow Graph

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Planning Is Constrained Graph Construction

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Model-Facing Projections

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Generated Runtime Artifacts

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Runtime Execution

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

An Unsafe and Corrected Compiler

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security and Supply-Chain Considerations

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Operational Implications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Generate

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: The Specification Passed but the Deployment Drifted

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 10 – Deterministic and Probabilistic Boundaries

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Five Responsibility Zones

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

What the Model Should Interpret

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

What Deterministic Code Should Validate

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

What Policy Engines Should Authorize

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

What Transaction Services Should Execute

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

What Humans Should Approve

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Deterministic Shell

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Constrained Decoding and Schema Enforcement

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Confidence Thresholds

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

False Determinism

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

An Unsafe and Corrected Decision Path

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Latency and Cost Trade-offs

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security and Operational Implications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Use a Model

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: JSON That Is Valid and Wrong

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 11 — Orchestration, Delegation, and Multi-Agent Systems

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Multi-Agent Is an Organizational Choice

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Supervisor and Specialists

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Delegation Contracts

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Task Ownership

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Authority Propagation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Shared State Without Shared Confusion

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Handoff Schemas

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Result Validation and Trust

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Coordination Versus Decomposition

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Circular Delegation and Termination

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Unsafe and Corrected Multi-Agent Design

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security and Privacy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Reliability and Operations

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Use Multiple Agents

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: When More Agents Create Less Accountability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 12 — Failure Semantics and Recovery

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Failure Is a Contract Outcome

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Transient and Permanent Failures

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Invalid Model Output

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Tool Timeout and Unknown Outcome

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Partial Execution

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Duplicate and Out-of-Order Results

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Dependency Failure

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Approval Timeout and Denial

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Conflicting Decisions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Backoff and Jitter

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Dead-Letter Handling

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Compensation and Escalation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Retry Versus Safe Replay

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Failure Budgets

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Unsafe and Corrected Recovery

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security and Privacy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Operational Implications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Recover Automatically

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: Recovery Without Wishful Rollback

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Part IV – Verification and Quality

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 13 – Testing Agent Specifications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

A Layered Test Strategy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Schema Tests

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Parser and Normalization Tests

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Semantic Tests

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Contract Tests for Tools

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

State-Transition Tests

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Invariant Tests

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Property-Based Testing

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Model-Based Testing

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Simulation and Failure Injection

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Adversarial Tests

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Golden Datasets and Regression Suites

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Mutation Testing

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Test Doubles and Deterministic Models

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Stochastic Test Strategy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security and Operations

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Use an End-to-End Model Test

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: A Test Suite That Passed the Wrong System

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 14 — Evaluating Behavior, Not Just Answers

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Evaluation Begins With a Behavioral Record

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Separate Hard Constraints From Quality

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Response Quality

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Task Completion

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Tool Selection and Argument Correctness

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Policy Compliance and Execution Order

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Escalation Quality

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Evidence Use and Source Attribution

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Latency, Cost, and Reliability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Behavioral Scorecards

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Offline Evaluation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Online Evaluation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Evaluation Leakage

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Evaluator Disagreement

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Unsafe and Corrected Evaluation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security and Privacy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Use a Single Metric

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: The 92 Percent Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 15 — Guardrails and Runtime Enforcement

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

A Guardrail Must Control a Boundary

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Input Validation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Output Validation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Domain Validation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Policy and Authorization

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Rate Limits and Budgets

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Sandboxing

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Content Controls

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Data-Loss Prevention

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Approval Gates

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Kill Switches and Capability Rollback

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Layered Guardrails

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Unsafe and Corrected Engineering Guardrails

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Guardrail Composition and Failure

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security and Privacy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Add Another Guardrail

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: The Guardrail That Lived in the Prompt

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 16 – Observability, Auditability, and Replay

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Three Evidence Products

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Traces and Spans

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Execution Timelines

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Decision Records

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Version Identity

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Tool Calls and Policy Decisions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Provenance

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Token and Cost Accounting

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Audit Logs

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Redaction

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Correlation Identifiers

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Safe Replay

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Forensic Investigation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Unsafe and Corrected Observability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Privacy Implications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Store a Transcript

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: Reconstructing a Duplicate Effect

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Part V — Production Patterns

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 17 — Customer-Support Triage Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Scope and Architecture

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

AgentSpec Boundary

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Data Model

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

State Machine

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Tool Contracts

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Policies and Privacy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Classification and Deterministic Rules

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Testing Strategy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Walkthrough: The Suspicious-Login Ticket

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Decision Table

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Data-Threat Review

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Observability and Audit

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Failure Analysis

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Unsafe Design

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Corrected Design and Trade-offs

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Use This Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: Mixed Intent Under Degraded Dependencies

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 18 — Software-Engineering Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Scope and Architecture

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

AgentSpec Boundary

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Repository Permissions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Sandboxing and Command Contracts

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Planning Contract

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Patch and Artifact Provenance

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Test Gates

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Pull-Request Preparation and Human Review

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Walkthrough: Protected Workflow Drift

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

State and Failure Paths

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Evaluation and Tests

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Observability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Trade-offs

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Use This Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: A Correct Patch from an Unsafe Process

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 19 — Financial-Operations Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Scope and Architecture

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

AgentSpec Boundary

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Authoritative Data and Immutable Evidence

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Threshold Policy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Separation of Duties

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Review Package

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

State and Human Authorization

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Idempotency and Duplicate Containment

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Compensation and Reversal

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Data Privacy and Explainability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Walkthrough: Threshold Drift

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Tests and Evaluation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Observability and Audit

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Failure Containment

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Trade-offs and When Not to Use

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: The Threshold at the Currency Cutoff

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 20 — Multi-Agent Incident Response

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Roles and Authority

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Architecture

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Supervisor AgentSpec

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Immutable Incident Timeline

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Evidence Collection

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Conflicting Hypotheses

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Time Bounds and Budgets

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Blast-Radius Control

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Remediation State Machine

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Communications Contract

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Walkthrough: The Wrong-Region Hypothesis

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Failure and Recovery

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Testing and Evaluation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Observability and Post-Incident Replay

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Trade-offs and When Not to Use

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: Approval Invalidated by Late Telemetry

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 21 — Portability, Evolution, and Governance

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Model Portability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Framework Portability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Specification Versioning

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Schema Evolution

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Migration

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Deprecation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Governance as an Engineering System

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Review Processes

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Agent Catalogs

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Organizational Maturity

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production-Readiness Gates

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Unsafe and Corrected Evolution

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security and Operational Implications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When Not to Pursue Portability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Design Review: A Portable Specification on an Incomplete Framework

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Production Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Key Takeaways

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Exercises

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

From Agent Demos to Engineered Systems

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Discipline of Agent Specification Engineering

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Central Separations

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

A Practical Adoption Sequence

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

1. Inventory purpose and authority

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

2. Define ports and completion

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

3. Contract every tool

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

4. Separate information planes

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

5. Enforce state and policy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

6. Engineer failure

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

7. Test behavior

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

8. Make evidence operational

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

9. Govern evolution

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

A Ninety-Day Adoption Plan

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Days 1–15: Establish the Boundary

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Days 16–35: Make the Contract Executable

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Days 36–60: Attack the Contract

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Days 61–75: Observe Without Expanding Authority

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Days 76–90: Controlled Capability and Governance

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Artifacts That Prove Progress

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

When to Stop the Adoption

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The First Design Review Agenda

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

What AgentSpec Contributes

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Standard for “Reliable”

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

The Engineering Model to Carry Forward

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Appendix A — AgentSpec 0.1 Language Reference

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Document and Serialization

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Conformance Levels

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Top-Level Inventory

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

metadata

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

purpose and vocabulary

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Data Ports: inputs and outputs

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

objectives

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

capabilities

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

tools

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Predicates

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

contract

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

state

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

policies

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

memory

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

context

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

data

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

budgets

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

resilience

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

observability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

audit

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

evaluation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

portability

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Validation Order

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Minimum Runtime Events

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Compatibility and Migration

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Appendix B — AgentSpec JSON Schema

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Dialect and Identity

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Closed Top-Level Structure

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Reusable Definitions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Required Versus Optional Fields

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Enumerations

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Bounds

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Arrays and Uniqueness

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Embedded Port and Tool Schemas

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Schema Validation in the Reference Engine

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Versioning the Schema

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Using the Schema in CI

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Complete AgentSpec 0.1 Schema

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Appendix C — Threat-Modeling Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

1. System and Purpose

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

2. Assets

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

3. Entry Points and Trust Boundaries

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

4. Identity and Delegation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

5. Capability and Tool Authority

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

6. Prompt Injection and Untrusted Content

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

7. Data Privacy and Isolation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

8. Context and Memory Integrity

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

9. Model and Provider Boundary

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

10. State and Workflow Integrity

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

11. Side Effects and Distributed Failure

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

12. Sandbox and Execution Environment

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

13. Human Approval and Social Engineering

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

14. Output and External Communication

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

15. Observability, Audit, and Replay

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

16. Supply Chain and Control Plane

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

17. Availability and Abuse

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

18. Evaluation and Detection

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

19. Residual Risk and Decision

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Threat-Model Review Record

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Appendix D — Production-Readiness Checklist

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Release Header

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 1 — Purpose, Ownership, and Scope

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 2 — Specification and Build Integrity

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 3 — Identity, Authorization, and Secrets

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 4 — Data Governance and Privacy

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 5 – Tooling and Side-Effect Safety

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 6 – State, Concurrency, and Recovery

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 7 – Model and Context Behavior

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 8 – Sandbox, Network, and Supply Chain

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 9 – Verification and Evaluation

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 10 – Observability and Audit

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 11 – Operations and Incident Response

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Gate 12 – Rollout and Change Control

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Release Decision

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Fast Red-Flag Review

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Appendix E — Evaluation Scorecards

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Scorecard Header

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Universal Hard Gates

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Behavioral Rating Scale

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Core Behavioral Dimensions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Outcome Correctness

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Evidence Grounding

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Planning and Tool Efficiency

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Argument and Schema Quality

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Escalation Quality

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Failure and Recovery Behavior

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Communication Quality

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Weighting Template

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Risk Slices

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Customer-Support Triage Scorecard

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Software-Engineering Agent Scorecard

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Financial-Operations Agent Scorecard

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Multi-Agent Incident Scorecard

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Online Evaluation and Stop Conditions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Adjudication Protocol

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Completed Scorecard Record

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Appendix F — Exercise Solutions and Discussion Notes

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 1 — The Agent Reliability Problem

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 2 — From Prompts to Behavioral Contracts

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 3 — Anatomy of an Agent Specification

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 4 — Goals, Outcomes, and Completion Conditions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 5 — Tool Contracts and Side Effects

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 6 – State, Memory, and Context

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 7 – Decisions and State Transitions

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 8 – Policies, Permissions, and Authority

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 9 – Compiling Specifications into Workflows

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 10 – Deterministic and Probabilistic Boundaries

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 11 – Orchestration, Delegation, and Multi-Agent Systems

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 12 – Failure Semantics and Recovery

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 13 – Testing Agent Specifications

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 14 – Evaluating Behavior, Not Just Answers

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 15 – Guardrails and Runtime Enforcement

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 16 – Observability, Auditability, and Replay

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 17 – Customer-Support Triage Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 18 – Software-Engineering Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 19 – Financial-Operations Agent

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 20 – Multi-Agent Incident Response

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Chapter 21 – Portability, Evolution, and Governance

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Appendix G — Glossary

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Appendix H – Bibliography and Further Reading

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Contracts, Formal Reasoning, and State

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Data Formats, APIs, and Events

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Security, Privacy, and Risk

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Observability, Reliability, and Supply Chain

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Language Models, Reasoning, and Tools

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Suggested Reading Paths

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.

Source Maintenance Note

This content is not available in the sample book. The book can be purchased on Leanpub at <https://leanpub.com/agentspecificationengineering>.