DOCUMENT360

Webinar On

Embedding Critical Thinking into AI-Assisted Documentation

Three Reusable Frameworks for Clarity, Coherence, and Responsibility

Speaker

Pola Drakopoulou

Technical / UX Writer

Mitel

DOCUMENT360

The Agenda

- Failure patterns in AI-assisted documentation
- How the technical writer’s role is changing
- Three reusable thinking frameworks
- Two real workflows-examples
- Outcome, trade-offs, and honest limits

Three Patterns of AI Failure

Without systematic guidance,

AI output may overgeneralize across contexts

AI doesn't automatically address:

Ethics, security, privacy, or accessibility.

AI makes unstated assumptions:

About what users know, who your audience is, and what context you're in

The Structural Bias of LLMs

Based on a case study using Claude & GPT-4o Mini

- Computational preference to keep what is already embedded
- Replacing output implies original choice was wrong
- Easier to defend known patterns with less uncertainty
- "I created X" equals "X meets requirements"

Documentation Quality in an AI world

76% of writers regularly use AI, but only 44% have guidelines in place.

Source: The State of Docs Report 2026 (stateofdocs.com/2026/ai-and-documentation-creation).

The risk:

- Approving AI-generated content because it sounds right.

The challenge:
Most teams do not yet have:

- Built-in AI guardrails
- Traceable reasoning
- Enforced quality controls

How do we systematically control quality, judgment, and risk today, while AI is part of the workflow?

DOCUMENT360

Writer as Architect

Documentation quality is no longer a writing problem. **It is a systems design problem.**

The technical writer role is shifting:

producing content → designing, and governing the system that produces content reliably.

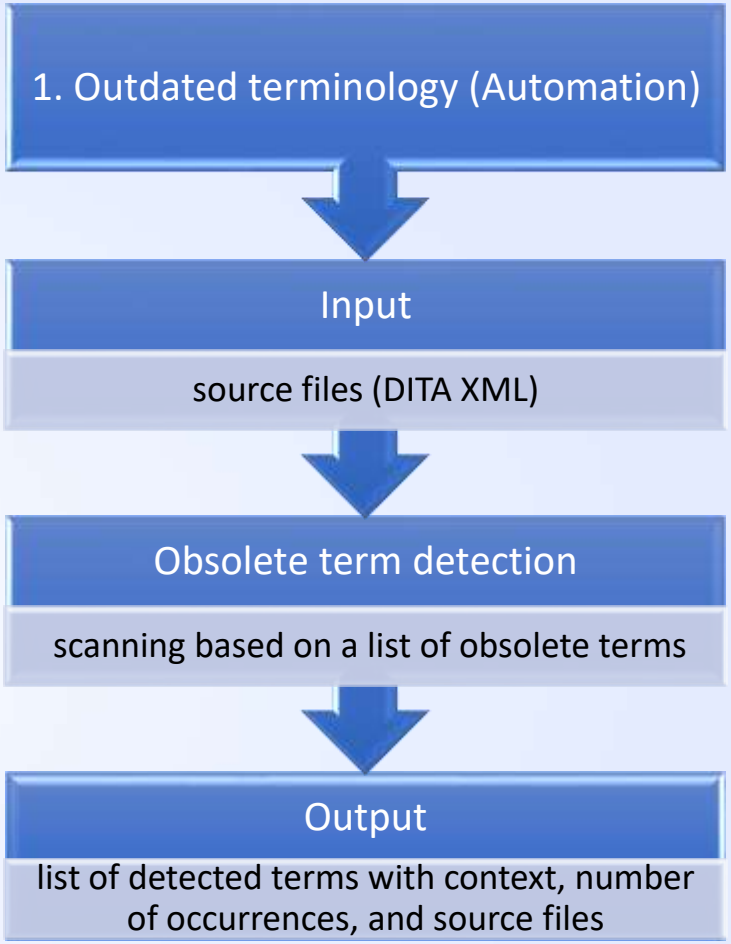
DOCUMENT360

The Automation Layer

The problem

Improving a legacy **documentation set** with:

1. Outdated terminology
2. Chapters with inconsistent quality and gaps



The Thinking Layer

2. Chapter quality review			
What is the goal?	**What assumptions are we making?**	**Is the level of detail right?**	**What could go wrong?**

From Questions to Frameworks

2. Chapter quality review

What is the goal?	What assumptions are we making?	Is the level of detail right?	What could go wrong?

1 **Golden Mean:**
What is the goal? Is the level of detail right?

2 **Socratic Method:**
Assumptions

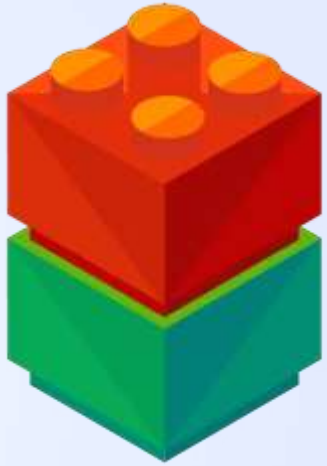
3 **Ethical Editor:**
Compliance, security risk, accessibility

What Are Frameworks?

Reusable control (‘thinking’) layers.

They encode:

- Rules
- Constraints
- Evaluation criteria



They are:

- Reusable
- Implementation-agnostic
- Portable

Why Control Layers Matter:
From Operating AI to Governing It

- Shape **mental models** on responsible, deliberate AI governance
- Operate → **orchestrate** AI systems
- From ‘using a tool’ → ‘interacting with a tool’
- Passive consumption → critical thinking, and **active sense-making**
- Can intentionally **introduce productive complexity** to force deeper, traceable reasoning
- **Ownership** of our decisions

Meet the Frameworks: Golden Mean

Purpose: Information calibration to the user's goal, low cognitive load

TOO LITTLE	GOLDEN MEAN	TOO MUCH
Confusing	*Clarity*	*Overwhelming*

Golden Mean Checklist

- ✓ Does the detail level match the importance of the task?
- ✓ Can I remove anything without losing clarity?
- ✓ Does every paragraph serve the user's goal?
- ✓ Are deep dives linked rather than inline?

Output: Flagged sections marked as over-verbose or under-specified.
Writer decides: cut, expand, or link out.

Meet the Frameworks: Socratic Questioning

Purpose: Challenge assumptions through sustained inquiry

Context	Process	Outcome

Socratic Questioning Checklist

Context

- What prior knowledge is assumed?
- What environment or setup is assumed?
- Are all prerequisites explicitly stated?

Process

- What if this step fails?
- Are there unstated dependencies?
- What happens between steps?

Outcome

- How does the user know it worked?
- What are realistic failure scenarios?
- Is a rollback or recovery path documented?

Output: Missing prerequisites, missing verification steps, etc.
Writer decides: confirm, revise, or escalate assumptions.

Meet the Frameworks: Ethical Editor

Purpose: Detect security, compliance, and accessibility issues

- Security
- Accessibility
- Privacy & Compliance
- Impact

Ethical Editor Checklist

Security & Safety

- Security warnings for risky operations?
- Dangerous defaults flagged?

Accessibility

- Screen reader compatible?
- No visual-only instructions (color, position)?

Privacy & Compliance

- Personal data collection disclosed?
- User consent requirements documented?

Impact

- Irreversible actions clearly marked?
- Potential harm scenarios addressed?

Output: Risk level (Low / Medium / High)
Writer decides:
Medium → targeted review
High → escalation and peer review

Workflow 1: Frameworks in an Agent

Input: PDF, text (i.e., from Jira description, Confluence page)

User selects mode (Review/Edit/Draft)

↓

Agent collects context: goal, audience, document type, risk level

↓

Applies control layers

↓

Outputs:

- Issues and gaps
- Risk level per finding
- Recommendations

Workflow 1: Summary Report

Administrator guide
Total issues: 14, High-risk: 2

1

Golden Mean (5 issues):

Missing prerequisites · Concepts mixed with steps · Dense command blocks · No task grouping · No step outcomes

2

Socratic Method (5 issues):

Unclear execution context · Prior knowledge assumed · No validation steps · No failure scenarios · No rollback guidance

3

Ethical Editor (2 critical, *4 issues in total*):

HIGH: System accounts may have insecure permissions or configurations that could allow unauthorized access · Some security settings may weaken system protection, but the report does not specify which ones

Workflow 2: A Docs-as-Code Workflow

Input: selected code snippet from dev repo

1. Run a generation script to produce initial artifacts:

File	Purpose
`impact.md`	Documentation impact analysis derived from code
`draft.md`	Human-readable draft from code
`review.md`	Framework review populated with findings
`concept.dita`	DITA concept topic candidate (original + ‘refined’)
`task.dita`	DITA task topic candidate (original + ‘refined’)

2. Rewrite concept/task outputs based on review
3. Review, validation with SMEs, and final editorial decisions (Non-automated)

Workflow 2

review.md

```
## Issues by Framework

### Golden Mean
- Issue: The generated draft may still need one short statement about when an administrator should choose enable vs disable.
  Severity: MEDIUM
  Recommendation: Add a decision-oriented sentence tied to a real admin scenario.

### Socratic Method
- Issue: The selected snippet does not fully show end-to-end save and validation behavior.
  Severity: MEDIUM
  Recommendation: Verify expected result and confirmation behavior from the full page workflow.

### Ethical Editor
- Issue: Permission-based disabling can be misread as a defect.
  Severity: MEDIUM
  Recommendation: Add explicit permission notes where controls can be visible but unavailable.
  Escalation: Confirm wording with SME/product owner if role behavior is customer-sensitive.

### Information Architecture
- Issue: Concept and procedure content are both present.
  Severity: LOW
  Recommendation: Keep both concept and task topics, then publish the one that best fits chapter intent.
```

Workflow 2: Summary Report

Administrator guide (code-derived draft)
Total issues: 3, High-risk issues: 0

1. **Golden Mean (1 issue):**
   Missing decision context for when an administrator should enable or disable the setting.

2. **Socratic Method (1 issue):**
   The selected snippet does not fully show end-to-end save and validation behavior.

3. **Ethical Editor (1 issue):**
   If documentation does not explain permission restrictions, users may think a disabled feature is broken.

Workflow 2

`task.dita`

Prerequisites:

- Main [redacted] write permission, [redacted] tab write permission, or [redacted] write permission.

`task.dita (refined)`

Prerequisites:

Ethical Editor: permission note added so unavailable controls are not misread as defects

Before you begin, confirm the following:

- The [redacted] feature flag is active in your system. If the section does not appear in the UI, this feature is not available in the current version or patch set.
- You hold at least one of the following permissions: [redacted] write permission, [redacted] tab write permission, or [redacted] write permission. Without one of these, the control will be visible but unavailable (grayed out).

Escalation (Ethical Editor): confirm exact permission names and customer-facing wording with SME/product owner

Design the Governance Layer

1. Define the problem and whether it needs a control layer or just automation

2. Identify the ‘thinking layers’ for your domain

- Frameworks complement existing style guides and review gates

3. Design the control flow for your domain

4. Automate execution

Who Decides

Layer	Owned by
Style guide / terminology	Editorial standards
Framework review	Writer and AI as assistant
Domain accuracy	SMEs
Publication approval	Existing gate

What Frameworks Are
(and What They Are Not)

- Supports structured reasoning
- Human-in-the-loop by design
- **Bi-directional utility**: Forward (generative constraints) & Backward (evaluative quality gates)

These frameworks **are not**:

- Autonomous correctness engines or text generators
- A replacement for domain expertise
- A system that approves, validates, or publishes content

Aligns with emerging best practices: 56% of teams still lack guidelines, and top practitioners report using semi-automation specifically to ensure the writer has the ultimate say.

DOCUMENT360

Source: The State of Docs Report 2026 (stateofdocs.com/2026/ai-and-documentation-creation).

What Changed

Early observations

BEFORE	AFTER
Reactive	Proactive
Issues found after release	Issues caught before release
Inconsistent review	Structured, consistent reasoning
Heavily senior-dependent review	Scalable across team

The Reality Check:
Trade-offs and Implementation Costs

What this approach costs

- Additional upfront review time
- Occasional false positives requiring review

Why the cost is justified

- Review effort is focused based on the risk level
- Issues caught earlier (also cheaper to fix)
- More consistent decision-making across reviewers
- Fewer hidden issues reach publication

Closing Thoughts

64% believe AI will be 'extremely impactful’

Source: The State of Docs Report 2026

Until outputs are fully traceable and explainable (and even then) **human** judgment, validation, and governance remain the most essential part of our work.

DOCUMENT360

Thank You!

Q & A

DOCUMENT360