

# DOCUMENT 2: THE AUDIT PROMPT

## VAULT OF AGES — ROMAN HISTORICAL ACCURACY

### EVALUATION SYSTEM

© 2026 NSH PUBLISHING, LLC — ALL RIGHTS RESERVED

---

#### HOW TO USE THIS DOCUMENT

Submit this prompt to Claude with up to 15 Midjourney-generated images attached. Claude will evaluate every image against the Roman Historical Consistency Ruleset and return a structured verdict, issue report, and showcase rating for each image.

**Structural layers** — Visual Anchoring Pass, Evidence Confirmation Pass, Negative Knowledge Gate, Rating Calibration Guide, and Structured Issue Handoff Format — are period-agnostic and do not change when adapting for other historical periods.

**The Roman Historical Consistency Ruleset** is the period-specific layer. When adapting for Viking Age, ancient Greece, feudal Japan, or other periods, replace this ruleset with period-appropriate constraints. Everything else remains identical.

---

#### THE PROMPT *(Submit to Claude with images attached)*

---

Attached are up to 15 AI-generated images intended to be historically accurate to ancient Rome (roughly 753 BC – 476 AD). Please examine every image and evaluate each one using the criteria and rules below.

Default evaluation standard is conservative 1st–3rd century CE Roman unless the image clearly requires another Roman-period context.

---

#### VISUAL ANCHORING PASS (MANDATORY)

Before performing any historical or logical evaluation, identify the primary visual elements present in each image.

List the observed elements briefly under the following headings:

Observed elements:

People:

Clothing / armor:

Weapons / equipment:

Architecture:

Environment:

Lighting and surface condition:

Only list elements that are clearly visible. Do not infer hidden or assumed objects. This pass ensures that all subsequent evaluation is based only on visible evidence.

---

## EVIDENCE CONFIRMATION PASS (ANTI-HALLUCINATION GUARD)

Before assigning any violation, confirm that the suspected issue is supported by clearly visible evidence.

### Procedure:

- Identify the specific object, feature, or detail that appears problematic
- Confirm that the element is clearly visible in the image
- Verify that the element directly violates the Roman Historical Consistency Ruleset
- If the element cannot be clearly observed or confidently identified, mark it as uncertain rather than assigning a violation

### Rules:

- Suspicion alone is not sufficient to assign a failure
  - If you cannot point to a specific visible object or feature, the issue must not be flagged
  - Every violation must correspond to visible evidence that can be described explicitly
- 

## NEGATIVE KNOWLEDGE GATE (FALSE POSITIVE PREVENTION)

Distinguish between:

- Visible violation
- Uncertain visibility
- Not visible / not assessable

### Rules:

- Do not flag an anachronism unless there is clear visible evidence
- Do not assume hidden details
- Do not infer incorrect equipment, materials, hairstyles, footwear, architecture, or symbols if they are cropped, shadowed, blurred, obstructed, or too small to inspect reliably
- Lack of visible confirmation is not the same as evidence of error
- Only assign REJECT when the violating feature is clearly visible and confidently identifiable ✖
- Use BORDERLINE for uncertainty only when it meaningfully affects historical evaluation ⚠

**Priority rule:** Visible evidence overrides suspicion. Suspicion without visible evidence must not produce rejection.

---

## ROMAN HISTORICAL CONSISTENCY RULESET (CANONICAL)

*(This is the period-specific layer. Replace with period-appropriate constraints when adapting for other historical settings.)*

All visible elements must be historically plausible for the Roman world within the conservative evaluation window (1st–3rd century CE) unless the image clearly depicts another Roman-period context.

### **Lighting**

- Open flame only: torches, oil lamps, braziers
- No electric lighting of any kind
- No glass-paneled lamp housings
- No pipe runs or conduit

### **Clothing**

- Roman garments only
- No zippers, elastic fabrics, or modern tailoring

### **Footwear**

- Roman caligae or bare feet only

### **Armor**

- Roman armor types only
- No medieval plate armor, enclosed visors, Gothic pauldrons, or articulated gauntlets

### **Weapons**

- Gladius or spatha only
- Swords worn on the hip, never on the back
- Shields must be Roman forms only (scutum or parma)

### **Architecture**

- Round Roman arches only
- No Gothic, medieval, Renaissance, Baroque, or modern structures
- No clear or flat window glass
- Primary materials: brick, stone, tufa – not white marble

### **Symbols**

- No crosses, Christian iconography, heraldry, fleur-de-lis, or coats of arms
- No Arabic numerals

### **Materials**

- No plastic, chrome finishes, synthetic fabrics, modern machined hardware, I-beams, rebar, or modern running bond brick

### **Hair**

- Roman male hairstyles only
- No modern fades, undercuts, or styled gel hair

---

## **EVALUATION CRITERIA**

Apply all six criteria to every image:

- 1. Historical Anachronism Check** Evaluate against the Roman Historical Consistency Ruleset. Flag any clearly visible element that belongs to a later historical period or the modern world. Apply the Negative Knowledge Gate — only flag violations that are clearly visible.
  - 2. Cultural and Equipment Consistency** Check whether clothing, armor, weapons, and equipment belong to the Roman world and are used in culturally plausible combinations. Flag Roman gear mixed with medieval or non-Roman equipment, incompatible equipment combinations, and shields, armor, or weapons used in implausible ways.
  - 3. Gladiator Accuracy** For gladiator images: historically plausible body type (substantial build, visible body fat over muscle — not lean modern athletic physique), correct weapon placement (sword on hip, not back), plausible gladiator equipment combinations, period-appropriate hairstyles.
  - 4. Architectural Accuracy** If architecture appears: Roman round arches only, no Gothic or medieval structural elements, no Renaissance or later styles, no modern building materials. Narrative battle damage is acceptable if the structure still appears Roman in origin.
  - 5. Social Role Consistency** Evaluate whether clothing, equipment, posture, and setting align with plausible Roman social roles. Flag slaves dressed as elites, gladiators in aristocratic clothing, senators in battlefield armor without narrative context, contradictory status signals.
  - 6. Showcase-Worthiness** Rate visual strength, composition, atmosphere, and usefulness as showcase artwork. Historical accuracy determines the verdict. The star rating evaluates visual impact.
- 

## AI ARTIFACT DETECTION

Flag if visible:

- Distorted hands or fingers, extra or fused fingers
- Duplicated limbs or missing joints
- Weapons or shields merging into bodies or clothing
- Objects intersecting incorrectly
- Impossible or inconsistent shadows
- Armor pieces floating without attachment
- Straps or belts that do not connect logically

Minor artifact → BORDERLINE ⚠ Severe artifact affecting the scene → REJECT ❌

---

## SURFACE AUTHENTICITY CHECK

Roman environments should appear used and lived-in, not pristine. Flag unrealistically perfect surfaces — mirror-polished armor, spotless marble, perfectly clean textiles, sterile studio-like environments.

Expected Roman surface characteristics: dust on arena sand, soot near lamps and torches, scuffed leather, patina on bronze or iron, weathered wood or rough masonry.

Minor cleanliness → BORDERLINE ⚠ Sterile or artificial environment → REJECT ❌

---

## RATING CALIBRATION GUIDE

Only KEEP and BORDERLINE images receive a rating. REJECT and ERROR images are not rated. When uncertain between two ratings, choose the lower score.

### ★★★★★ (5 stars)

Exceptional showcase image. Striking cinematic composition. Convincing historical atmosphere. No visible AI artifacts. Could function as cover art or promotional material.

### ★★★★ (4 stars)

Strong, high-quality image. Clear composition and strong focal subject. Convincing Roman atmosphere. Minor imperfections that don't distract. Suitable for catalog publication.

### ★★★ (3 stars)

Good but not exceptional. Historically plausible and visually readable. Composition may feel ordinary. Minor visual imperfections visible. Usable but not standout.

### ★★ (2 stars)

Weak composition or noticeable flaws. Scene understandable but visually unremarkable. Noticeable AI artifacts or realism issues. Historically plausible but visually mediocre.

### ★ (1 star)

Barely usable. Composition unclear or poorly framed. Major realism problems. Technically passes historical checks but visually ineffective.

---

## STRUCTURED ISSUE HANDOFF FORMAT (MANDATORY)

When an image receives BORDERLINE or REJECT, emit a structured issue report. Each detected problem is a separate Issue block.

**BORDERLINE verdicts must include a cited reason.** A BORDERLINE verdict with no stated reason is invalid. The reason must identify the specific uncertainty or marginal issue that prevented a KEEP verdict.

Issue N

Tag: [CANONICAL TAG FROM LIST BELOW]

Severity: [minor / moderate / major]

Visible evidence: [exactly what in the image triggered the failure]

Repair implication: [specific visual correction required]

---

## CANONICAL FAILURE TAG VOCABULARY

Use only these tags. Do not invent new tags. Do not modify tag names.

ANACHRONISM\_LIGHTING

ANACHRONISM\_CLOTHING

ANACHRONISM\_FOOTWEAR

ANACHRONISM\_ARMOR

ANACHRONISM\_WEAPON

ANACHRONISM\_ARCHITECTURE

ANACHRONISM\_SYMBOLS

ANACHRONISM\_MATERIALS

ANACHRONISM\_HAIRSTYLE

AI\_ARTIFACT\_ANATOMY

AI\_ARTIFACT\_OBJECT

AI\_ARTIFACT\_STRUCTURE

EQUIPMENT\_LOGIC\_ERROR

COMPOSITION\_ERROR

SURFACE\_AUTHENTICITY





SOCIAL\_STATUS\_ERROR

A tag may only be applied if the violating element is clearly visible in the image.

---

## OUTPUT FORMAT

For each image provide:

- The filename
- Verdict:  KEEP /  BORDERLINE /  REJECT /  ERROR
- ★ rating (1–5) for all KEEP and BORDERLINE images
- For BORDERLINE: full structured issue report using Issue blocks, plus a cited reason for the borderline classification

- For REJECT: full structured issue report using Issue blocks
- For ERROR: note the failure reason (image did not load, unreadable, corrupt, etc.); no Issue blocks required

After all images, provide a summary table in a fenced markdown block with filenames, verdicts, and ratings.

If an image fails to load or cannot be evaluated, assign ERROR, report the filename and failure reason, and move on. ERROR images receive no star rating.

---

© 2026 NSH Publishing, LLC. Free for personal and commercial adaptation. Redistribution permitted with attribution to L. M. Hawkes / [HawkesAdventures.com](http://HawkesAdventures.com). Full methodology documentation available at [HawkesAdventures.com](http://HawkesAdventures.com).