

# Step 6 — Review UI & Human-in-the-Loop Validation

## Purpose

Introduce a human review layer to validate, correct, or reject LLM-extracted investor rights before export, ensuring **accuracy, auditability, and trust**.

Step 6 transforms *validated model outputs* into *decision-grade legal data*.

---

## Why Step 6 Exists

LLMs can extract structured legal rights, but **cannot be the final authority** in legal or investment workflows.

Step 6 addresses three critical gaps:

- **Ambiguity** (multiple interpretations)
- **Conflicts** (TS vs SHA, amendments)
- **Accountability** (who approved what, and why)

This step turns the pipeline from *automation* into a **reviewable system**.

---

## Inputs

From **Step 5 — Validation & Aggregation**:

- `extraction.jsonl`
- One record per:
  - `company`
  - `docset` (SHA, TS, amendments)
  - `clause_family` (veto, liquidity, exit, anti-dilution)
- Each item includes:
  - Structured fields (clause type, trigger, effect, thresholds...)
  - Evidence quote ( $\leq 25$  words)
  - Provenance (file, page, article)

- Confidence score
  - Flags ( `ambiguous` , `conflicting` , `missing_evidence` )
- 

## Core UX Principles

### 1. Evidence-first

- Every right is reviewed alongside its exact legal quote and source.

### 2. Explicit human decision

- Each item must be:
  - Accepted
  - Accepted with edits
  - Rejected

### 3. No silent corrections

- Any human override is logged with reviewer, timestamp, and comment.

### 4. Fast triage

- Reviewers focus first on:
  - Pending items
  - Conflicts
  - Low-confidence outputs

---

## UI Overview (Streamlit)

### 1. Portfolio Review Table

One row per extracted right.

Key columns:

- Clause family & type
- Beneficiary
- Confidence
- Status chip (Pending / Accepted / Edited / Rejected)
- Signal chips (Conflict / Low confidence)

- Evidence reference

Filters:

- Pending only
  - Flagged only
  - Clause family
  - Confidence threshold
- 

## 2. Detail Review Panel

Displayed when selecting a row.

### Evidence

- Immutable quote
- File, page, article

### Editable structured fields

- Clause type
- Beneficiary
- Trigger
- Effect
- Thresholds

### Decision actions

- Accept as-is
  - Accept with edits
  - Reject (comment required)
- 

## 3. Conflict Resolution

When conflicting clauses are detected:

- All conflicting versions are shown side-by-side
- Reviewer must explicitly:
  - Choose one

- Edit into a consolidated version
- Reject

No automatic precedence is assumed.

---

## Human-in-the-Loop Logic

- **Auto-advance:** after each decision, the UI automatically selects the next pending item
  - **Keyboard-friendly:** optimized for rapid expert review
  - **No backend mutation:** original model outputs remain unchanged
- 

## Outputs

### 1. Reviewed Rights

`data/out-fake/review/<company>/reviewed_rights.jsonl`

Adds:

- `review_status`
- `reviewer`
- `review_timestamp`
- `review_comment`
- Edited structured fields (if applicable)

### 2. Audit Log

`data/out-fake/review/<company>/review_log.jsonl`

Every decision recorded:

- Before / after state
- Reviewer identity
- Timestamp
- Justification

### 3. Export

- CSV/XLSX

- Only **accepted or edited** items
  - Ready for:
    - Investment memos
    - Legal summaries
    - Downstream analytics
- 

## Why This Step Is a Differentiator

Most LLM pipelines stop at **confidence scores**.

This system delivers:

- A **decision interface**, not just extraction
- Full traceability (model → evidence → human approval)
- A workflow comparable to:
  - Legal ops tooling
  - Palantir-style review layers
  - Institutional investment processes