ENTERPRISE BRIEFING

● LIVE DATA PIPELINE

US FEDERAL PATENT · TRADEMARK · COPYRIGHT LITIGATION

PROOF OF CONCEPT — NOT A PRODUCTION REPLACEMENT

IP Litigation Data Automation & Review Pipeline

A working proof of concept demonstrating how Chrisian can programmatically acquire, normalize, enrich, and structure public IP dispute and litigation data, with human review integrated where judgment or verification is required.

Built to explore where these capabilities could support enterprise legal-data operations.

Author: Chrisian Data Architecture

Scope: U.S. Federal Courts + EPO Boards of Appeal + EUIPO Case Law + WIPO Arbitration & Mediation Center + Japan IP High Court

Sources: CourtListener/RECAP + EPO + EUIPO + WIPO + Japan IP High Court

Live Demo: courtlistener.chrisianinc.com

The Opportunity & Where We Fit

Leading commercial IP litigation databases operate at significant global scale, combining automated data extraction with expert analysis to transform source material into normalized, verified, curated and searchable IP case data.

Chrisian has built a working proof of concept demonstrating that we can support this same class of data operation: programmatically acquiring data from multiple public sources, extracting and normalizing it into a common structure, enriching the resulting records, and routing cases that require verification or judgment into structured human-review workflows.

We are not proposing to recreate or replace an existing commercial IP litigation database. The purpose of this demonstration is to show your team what Chrisian can now do and start a conversation about where these capabilities — data automation combined with scalable human review — could provide additional capacity, support existing workflows, or enable new data initiatives.

488

LIVE RECORDS INGESTED

Patent, trademark, copyright, design, domain & utility model

6

IP / CASE CATEGORIES

Patent, trademark, copyright, design, domain disputes & utility model

5

INDEPENDENT LIVE SOURCES

CourtListener/RECAP + EPO + EUIPO + WIPO + Japan IP High Court

0

FIXTURES / MOCK RECORDS

Every demonstrated record originates from a live source

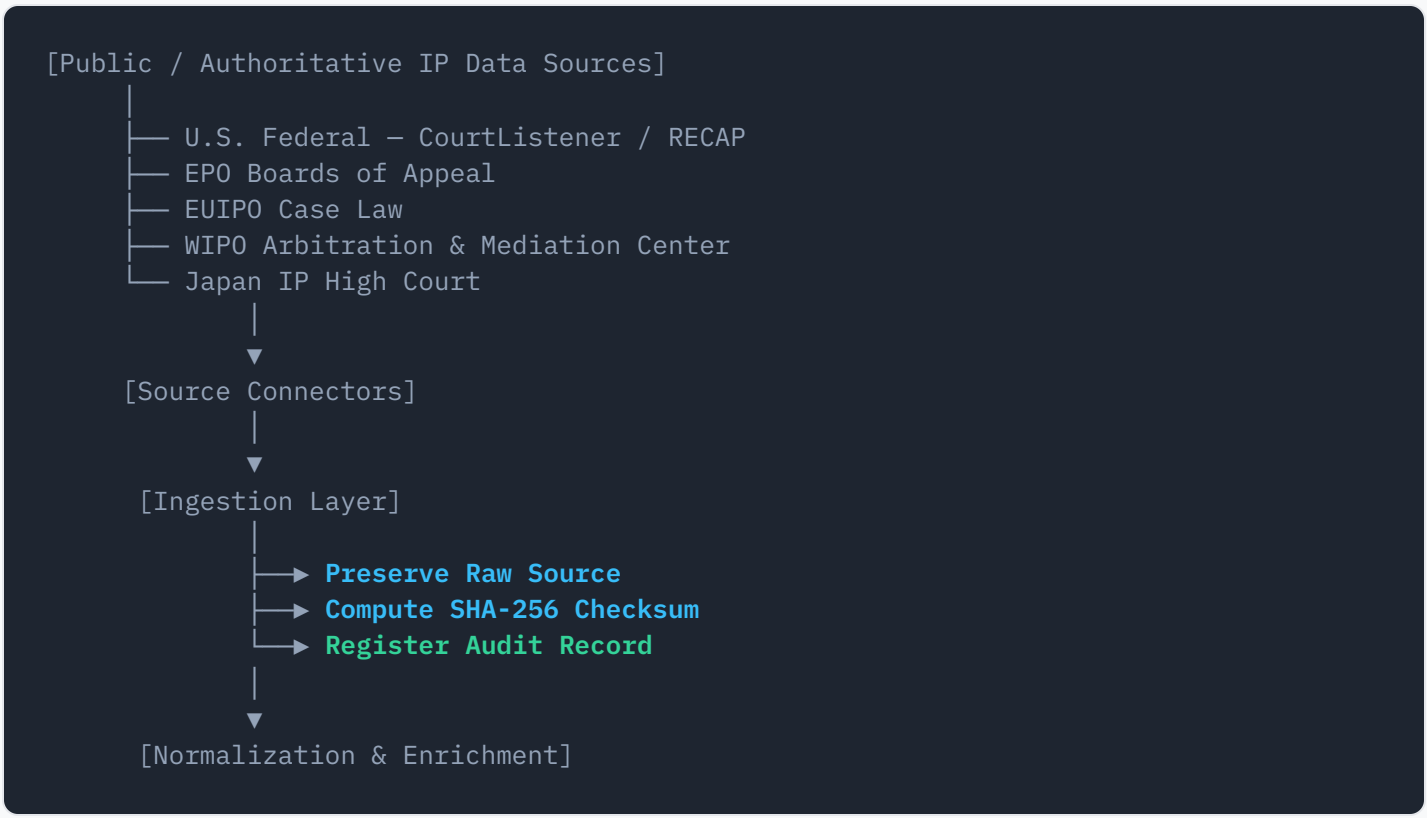
01 How We Acquire and Preserve Source Data

The pipeline is designed around source-specific connectors that programmatically acquire data from authoritative or established public data services. For U.S. federal litigation, the proof of concept connects to CourtListener and its RECAP collection of PACER data, operated by the nonprofit Free Law Project. CourtListener provides structured access to federal dockets, filings, parties, attorneys and related metadata.

The same connector-based architecture has been extended to additional IP data sources and jurisdictions, demonstrating that the pipeline is not dependent on a single source or data structure.

- **Programmatic acquisition:** source data is retrieved through dedicated connectors and transformed into a common processing pipeline.

- **Raw-source preservation:** each source response is preserved before transformation, creating an auditable record of the data received.
- **Integrity verification:** source snapshots are SHA-256 checksummed and registered so processed records can be traced back to their original input.
- **Extensible architecture:** new sources can be integrated through additional connectors while feeding the same normalization, enrichment and human-review framework.



02 How We Structure and Normalize the Data

Each source has its own data model, terminology and level of detail. The pipeline maps these disparate inputs into a common schema, creating consistent structured records that can be searched, enriched, reviewed and delivered regardless of the originating source.

Field group	What we capture	Status
Case identity	Jurisdiction, court/board, case number, case name, source identifier	`Programmatic`

Dates	Filing, decision, termination and other available procedural dates	`Programmatic`
Parties & representation	Parties, attorneys and law firms where available from the source	`Programmatic`
IP classification	Patent, trademark, copyright or design	`Programmatic`
IP rights referenced	Patent, trademark or design identifiers detected in source data/documents	`Programmatic + human review`
Procedural status	Pending, terminated, decided or other source-supported status	`Programmatic`
Outcome & legal issues	Disposition, prevailing party and applicable points of law	`Human review where required`
Source material	Linked filings, decisions, descriptions, dates, page counts and available text	`Programmatic`
Enrichment	Registration, ownership, chain-of-title and related data where available	`Programmatic + human review`

The result is a source-independent record structure: different jurisdictions and source systems can feed a common downstream workflow for search, enrichment, quality review and delivery.

Example of a normalized record generated from live source data:


```
{
  "case_name": "The American Psychiatric Association v. EmblemHealth, Inc.",
  "court": "District Court, S.D. New York",
  "docket_number": "1:25-cv-10783",
  "ip_type": "TRADEMARK",
  "date_filed": "2025-12-30",
  "disposition_status": "PENDING",
  "assigned_judge": "John George Koeltl",
  "parties": ["EmblemHealth Plan, Inc.", "The American Psychiatric Association"],
  "firms": ["Zuckerman Spaeder LLP", "Janove PLLC"],
  "documents": [ { "description": "NOTICE TO ATTORNEY REGARDING CIVIL...", "page_count":
}
```

03 What's Automated — and Where Human Review Adds Value

The proof of concept automates the repeatable stages of the data pipeline — source acquisition, raw-data preservation, normalization, classification, enrichment and search — without manual data entry between source retrieval and the resulting structured record.

Where the source data alone is insufficient to make a reliable determination, the system routes the record into a structured Human Review Console rather than forcing an automated conclusion. An analyst can review the underlying source material, confirm or correct the extracted information, and submit a decision that updates the structured record with a visible verification status and audit trail.

Outcome & legal-issue review. Determining disposition, prevailing party, grounds for a decision, or applicable points of law may require reading the underlying order or judgment rather than relying solely on docket metadata. The review console presents the source material alongside structured review fields, so the analyst's determination becomes part of the record.

AUTOMATION + HUMAN REVIEW

The objective is not to automate every decision. It is to automate repeatable work and create an efficient, auditable path for the work that requires human judgment.

Chrisian combines the technology to acquire, structure and enrich data with the production resources to operate structured review workflows at scale. This allows the same pipeline to support programmatic processing, exception handling, verification and analyst review within a single managed workflow.

This proof of concept is intended to demonstrate a capability that can be adapted to existing data operations or new initiatives based on the requirements of each workflow.

04 Human Review & Enrichment Workflows

The proof of concept includes eight additional review and enrichment workflows that demonstrate how programmatic processing and analyst review can work together. Each follows the same

principle: automate the repeatable work, surface exceptions or ambiguity, and route decisions requiring verification to a reviewer.

- **Entity De-duplication** — near-duplicate party and law-firm names are identified programmatically using similarity matching and presented to an analyst for confirmation before records are consolidated under a canonical entity name.
- **Federal Circuit Appeal Linking** — potential appellate relationships are identified through source-data searches and presented as candidate matches. An analyst verifies the relationship before an appeal is linked to the underlying district-court case.
- **EPO Translation Review** — German- and French-language EPO Board of Appeal decisions are ingested with their original source text and routed for English translation and review, allowing the resulting structured record to retain both source-language provenance and reviewed English content.
- **EUIPO Trademark Translation Review** — non-English EUIPO trademark decisions are acquired and routed through the same translation-review workflow. Where the source already provides structured information such as an outcome, that data is preserved rather than requiring the analyst to recreate it.
- **EUIPO Design Translation Review** — Registered Community Design decisions are processed as a distinct IP category, with design-specific review fields for legal issues such as novelty, individual character and informed-user comparison.
- **Machine-Drafted Translation** — the system can generate an initial English translation for analyst review and correction rather than requiring translation to begin from a blank field. The translation layer is provider-independent, allowing different translation services to be incorporated based on workflow requirements.
- **USPTO Patent Enrichment** — patent identifiers can trigger additional acquisition of USPTO registration and assignment data, including application information and chain-of-title history. The underlying history is preserved for analyst review rather than reducing potentially complex ownership information to an unsupported conclusion.
- **USPTO Trademark Enrichment** — trademark identifiers can similarly be enriched with USPTO registration data, including mark text, status, owner of record and goods/services classifications.

WHAT THIS DEMONSTRATES

Together, these workflows demonstrate Chrisian's ability to combine data automation with structured human review across multiple sources, jurisdictions, languages and IP categories.

The same architecture can support different levels of human involvement depending on the source and task: from simple verification of programmatically generated data to document review, translation QA, entity resolution and legal-data classification.

This proof of concept is designed to demonstrate the underlying capability rather than prescribe a particular client workflow. The specific balance between automation and human review can be adapted to the requirements of the data operation being supported.

05 Extending the Pipeline to New Sources & Jurisdictions

After establishing the workflow across U.S. federal courts, EPO and EUIPO, we extended the proof of concept to two substantially different sources: WIPO domain-name dispute decisions and Japan's IP High Court.

The objective was to test whether the underlying architecture could accommodate new jurisdictions, languages, document formats and types of IP proceedings without redesigning the normalization and human-review framework.

- **WIPO Domain-Name Disputes** — the pipeline now acquires decisions from the World Intellectual Property Organization's Arbitration and Mediation Center, which administers domain-name dispute-resolution proceedings including UDRP cases. 42 live decisions have been ingested across English, French and Spanish. Because these proceedings are dispute-resolution matters rather than court litigation, the pipeline models them as a distinct case type while retaining the same acquisition, normalization, translation and review framework used elsewhere in the proof of concept.
- **Japan IP High Court** — the pipeline has also been extended to decisions from Japan's IP High Court. 24 live decisions have been ingested across trademark, design and utility-model matters. Unlike the other sources in the proof of concept, the underlying decisions require PDF text extraction before their contents can enter the normalization pipeline. Japanese-language content can then move through the same machine-drafted translation and analyst-review workflow already used for European source material.

WHAT THIS DEMONSTRATES

Adding WIPO and Japan required source-specific connectors, but not a redesign of the underlying processing and review architecture.

The proof of concept now spans five live source systems, multiple jurisdictions and institutions, four languages, multiple document formats, and a growing range of IP and dispute types — all feeding a common framework for normalization, enrichment, translation, human review and delivery.

This demonstrates an architecture that can be extended to new data sources while preserving a consistent downstream workflow.

06 Search, API & Structured Delivery

The normalized data is available through both a searchable demonstration interface and a queryable REST API, allowing records to be explored interactively or accessed programmatically:

```
GET /api/v1/cases/search?ipType=PATENT&court=txed&limit=10
GET /api/v1/stats
```

Because every record lives in one common schema, batch export formats such as JSONL could be added for bulk delivery or downstream processing without restructuring the pipeline — not yet built in this proof of concept, but a natural next step.

Because acquisition, normalization, enrichment and human review all feed the same underlying data model, verified information is reflected consistently across search and API responses.

07 Where Chrisian Could Support Your Organization

This proof of concept is intended as a capability demonstration rather than a proposed replacement for an existing product or workflow.

It demonstrates Chrisian's ability to combine data automation with scalable human review — acquiring information from disparate sources, normalizing and enriching it into structured data, routing exceptions and judgment-based tasks for analyst review, and delivering the resulting data through searchable or programmatic interfaces.

We would welcome the opportunity to explore where these capabilities could support your organization — whether by adding capacity to an existing data operation, taking on a defined acquisition, enrichment or review workflow, or helping develop a new data initiative.

Inspect the System Live

The live demonstration portal provides access to records ingested from the source systems shown throughout this proof of concept, along with normalized data, source provenance, enrichment workflows and the Human Review Console.

Open Live Demo Portal →

© 2026 Chrisian, Inc. · Data Acquisition & Systems Architecture

Document: `docs/ip-litigation-pipeline-capability-brief.html`