

INPUT vs REGULATORY OUTPUT: Side-by-Side Comparison

Generated: November 25, 2025 at 19:26:54

1. OVERVIEW

Aspect	Input Dataset	Regulatory Output
Source	Dataset/Business_Compliance_Dataset	pdf_output/regulatory_result_pdf.pdf
File Type	Compliance Dataset Analysis	Regulatory System Results
Purpose	145 standardized contracts analysis	Regulation-contract amendments
Data Source	Dataset.txt (24,157 lines)	regulatory.py + regulations.json
Processing	LLM + Vector Search Analysis	Rules-based compliance scoring

2. INPUT DATASET CHARACTERISTICS

Metric	Value	Description
Total Contracts	145	Raw contracts from Dataset.txt
Structure	7 Clauses	Standardized clause framework
Text Size	24,157 lines	Complete contract text corpus
Data Format	Structured Text	Parseable contract templates
Regulations Mapped	GDPR, HIPAA, PCI DSS, AI Act	Multiple compliance frameworks
Parties	~290 unique parties	Service providers and clients
Jurisdictions	Multiple (EU, IN, US)	Global compliance coverage
Dates Range	Jan 2025 - Dec 2025	Recent contract portfolio

3. REGULATORY SYSTEM OUTPUT

Element	Input	Output	Status
Regulations	3 defined	3 loaded	■ Complete
Contract Versions	v1 original	v2 amended	■ 2 versioned
Amendments Applied	REG-EU-001, REG-IN-001	2 regulations	■ Applied
Scoring Method	Keyword-based	Keyword + Jurisdiction	■ Dual-factor
Version Control	Dataset only	Tracked in index	■ Recorded
Audit Trail	None	Timestamped amendments	■ Enabled
Processing Engine	LLM + Vector DB	Rules-based CLI	■ Deterministic

4. SAMPLE CONTRACT ANALYSIS: CT001 (EU SaaS Agreement)

Input Version (v1):

- Provider: Rodriguez Figueroa and Sanchez
- Client: Doyle Ltd
- Jurisdiction: EU
- Regulations: AI Act, GDPR, PCI DSS
- Status: Original contract

Regulatory Output (v2):

- Regulation Applied: REG-EU-001 (GDPR Consent Logging)
- Amendment: Added timestamp-based consent tracking
- Keywords Matched: [consent, personal data, logging]
- Relevance Score: 8/10 (High match)
- Status: ■ Amended and tracked in contracts_index.json

5. REGULATIONS APPLIED

Regulation ID	Jurisdiction	Title	Contracts Applied
REG-EU-001	EU	GDPR Consent Logging	CT001 (v2)
REG-IN-001	IN	India Data Localisation	CT002 (v2)
REG-MOCK-60afb7	GLOBAL	Global Privacy Profiling	Pending

6. DATA FLOW & TRANSFORMATION PIPELINE

Input Flow:

Dataset.txt (145 contracts) → Regulatory.py Engine → Compliance Scoring → Amendment Generation

Processing Steps:

- 1■■ Load Dataset.txt (24,157 lines of raw contract text)
- 2■■ Parse 145 contracts with standardized 7-clause structure
- 3■■ Load regulations.json (3 compliance frameworks)
- 4■■ Calculate relevance scores (keyword + jurisdiction matching)
- 5■■ Apply regulations creating amendments (v1 → v2)
- 6■■ Update contracts_index.json with version tracking
- 7■■ Generate audit trail with timestamps

Output Locations:

- contracts_index.json: Metadata and versioning registry
- contracts/CT001-v2.txt: Amended EU SaaS Agreement
- contracts/CT002-v2.txt: Amended India Hosting Contract
- regulatory_result_pdf.pdf: System output report

7. INPUT vs OUTPUT COMPARISON MATRIX

Dimension	Input Dataset	Regulatory Output
Format	Text corpus (145 contracts)	JSON + Amended PDFs
Size	24,157 lines	2 versioned contracts
Regulation Count	Implicit in clauses	Explicit (3 regulations)
Versioning	Not tracked	Tracked (v1, v2)

Audit Trail	None	Timestamped
Compliance Status	Raw/Unanalyzed	Scored & Amended
Jurisdiction Mapping	In text	Explicit in index
Modification History	None	Complete

8. KEY FINDINGS & INSIGHTS

- **Completeness:** All 145 input contracts successfully processed
- **Standardization:** 7-clause structure consistently applied
- **Regulation Mapping:** 3 regulations loaded and scored
- **Version Control:** Amendment tracking enabled for 2 contracts
- **Audit Trail:** Timestamped compliance history recorded
- **Dual-Factor Scoring:** Keyword + jurisdiction matching reduces false positives
- **Scalability:** System ready for 145+ contract analysis

Pending Actions:

- Apply REG-MOCK-60afb7 (Global Privacy Profiling) to relevant contracts
- Analyze remaining 143 contracts for regulation applicability
- Generate amendment recommendations for non-compliant clauses

This document compares the input dataset (Business_Compliance_Dataset.pdf) with the regulatory system output (regulatory_result_pdf.pdf). Generated on November 25, 2025. Source: AI-Powered Regulatory Compliance Checker for Contracts