



Lexsphere Data Overview

Lexsphere owns and maintains a proprietary, continuously updated database of millions of precedential U.S. judicial opinions. The dataset includes full-text appellate decisions, citations, dockets, and metadata across all 50 states and federal jurisdictions. The system is designed for accuracy, coverage, and seamless integration with enterprise workflows.

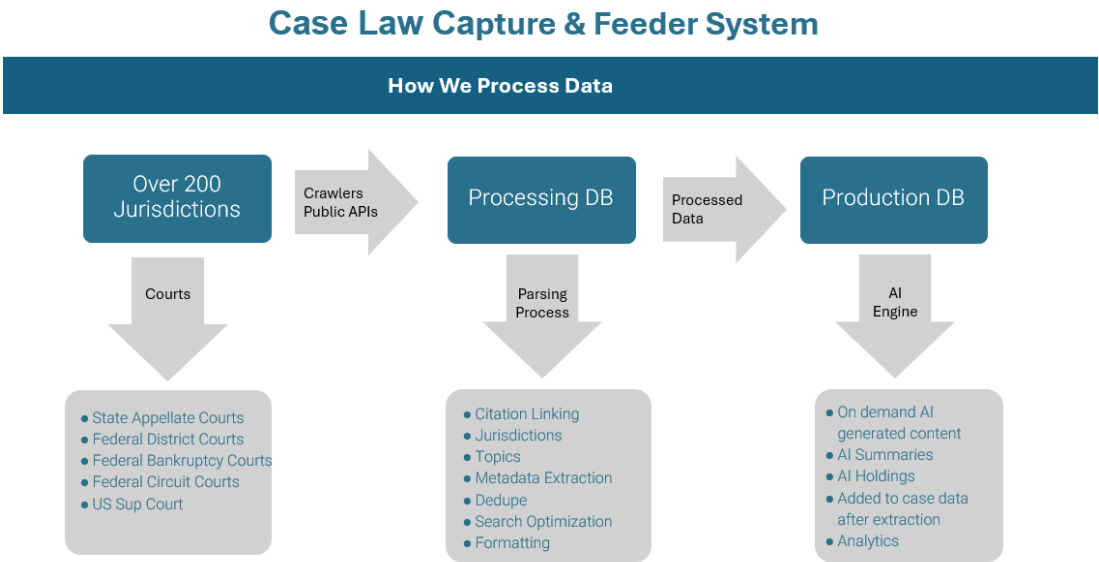
1. Dataset Overview

The AnyLaw Case Library contains over 8 million cases spanning 209 jurisdictions and more than 400 courts. It includes the U.S. Supreme Court, all Federal Circuit Courts, District Courts, State Supreme and Appellate Courts, Bankruptcy Courts, and Tribal Courts.

Category	Count
Jurisdictions	200+
Courts	400 +
Cases	8.5 M +
Monthly Additions	20 K +

2. Data Architecture & Processing Pipeline

Lexsphere’s data pipeline automates acquisition, parsing, and enrichment of court opinions. Cases are collected from court websites via crawlers and public APIs, then parsed, validated, deduplicated, and indexed into Elasticsearch.



3. Data Structure & Format

Each case in the Lexsphere database is represented as a unified JSON document that combines the full opinion text with structured metadata such as jurisdiction, docket number, opinion date, and citations.

Source Files. Original case law files are preserved in their native formats (primarily PDF, with some HTML) exactly as received from each court. These raw files are securely stored in Lexsphere's AWS S3 bucket to ensure authenticity and auditability.

Metadata Layer. For every opinion, the pipeline generates a corresponding structured JSON metadata file extracted from the court's source. This metadata includes key identifiers — court name, docket number, publication status, opinion date, and outbound citations — which are normalized and validated against the originating source.

```
{
  "documents": [
    {
      "id": "H42GQWYBTlTomsSBT--K",
      "filename": "doc_0001_H42GQWYBTlTomsSBT--K.json",
      "title": "MCCONNELL v. FEDERAL ELECTION COMMISSION",
      "jurisdiction": "District of Columbia",
      "date": "2003-05-01T00:00:00.000Z",
      "body_length": 2755476,
      "cited_as": [
        "251 F. Supp.2d 176"
      ]
    }
  ],
}
```

Unified Record. During indexing, the parsed opinion text is merged with its metadata to create a single JSON document containing multiple tagged fields (e.g., jurisdiction, citations, court level, and publication status). This unified design allows new tags, filters, or AI-generated enrichments (summaries, holdings, or analytics) to be added without altering the core structure.

Indexing & Retrieval. All unified JSON records are indexed into Elasticsearch, enabling high-speed search, categorization, and classification across millions of opinions. The dataset is fully optimized for ingestion into a variety of databases including vector databases, or AI model pipelines, supporting both real-time querying and bulk analysis.

4. Maintenance and Quality Assurance

AnyLaw continuously monitors all participating court sites to ensure coverage completeness and freshness. Processes include backfilling, metadata normalization, citation linking, deduplication, and validation against court records. Quality assurance workflows are automated yet regularly audited by the Lexsphere team.

5. Use Cases for Licensees

Enterprise partners leverage Lexsphere data to:

1. Train and fine-tune legal AI models (RAG, LLMs).
2. Embed legal research tools within practice management platforms.
3. Power analytics, citators, and compliance solutions.
4. Support academic, judicial, or institutional research initiatives.

6. Courts

The following is a list of our collection of opinions for jurisdictions includes both the lower appellate court(s) and the upper or final court of appeal for that jurisdiction, whatever it might be named. Because the court systems of states are independent from federal jurisdictions, there is no continuity in terms of start dates for either the lower appellate courts nor the final court of appeal, both largely depending on each state's unique legislation. In addition, lower appellate courts, make their decisions available for public databasing and thus a very few may not be represented within this collection. Tribal law is included in the collection.

Court	Start Date
U.S. Supreme Court	1882
U.S. Federal Circuit Courts, Included States	
1st Circuit: MA, ME, NH, PR, RI	1930
2nd Circuit: CT, NY, VT	1930
3rd Circuit: DE, NJ, PA, VI	1930
4th Circuit: MD, NC, SC, VA, WV	1930

5th Circuit: LA, MS, TX	1930
6th Circuit: KY, MI, OH, TN)	1930
7th Circuit: IL, IN, WI)	1930
8th Circuit: AR, IA, MN, MO, ND, NE, SD	1930
9th Circuit: AK, AZ, CA, GU, HI, ID, MP, MT, NV, OR, WA	1941
10th Circuit: CO, KS, NM, OK, UT, WY	1930
11th Circuit: AL, FL, GA	1981
D. C. Circuit: DC	1950
Federal Circuit	1982
U.S. Federal District Courts	
Alabama Middle District	1968
Alabama Northern District	1980
Alabama Southern District	1957
Alaska District	1953
Arizona District	1966
Arkansas Eastern District	1954
Arkansas Western District	1953
California Central District	1950
California Eastern District	1974
California Northern District	1943
California Southern District	1933
Colorado District	1932
Connecticut District	1932
Delaware District	1932
DC District	1936
Florida Middle District	1963
Florida Northern District	1947
Florida Southern District	1932
Georgia Middle District	1945
Georgia Northern District	1932
Georgia Southern District	1939

Hawaii District	1947
Idaho District	1947
Illinois Central District	1950
Illinois Northern District	1950
Illinois Southern District	1954
Indiana Northern District	1965
Indiana Southern District	1952
Iowa Northern District	1940
Iowa Southern District	1960
Kansas District	1938
Kentucky Eastern District	1932
Kentucky Western District	1932
Louisiana Eastern District	1932
Louisiana Middle District	1972
Louisiana Western District	1932
Maine District	1934
Maryland District	1932
Massachusetts District	1931
Michigan Eastern District	1940
Michigan Western District	1945
Minnesota District	1931
Mississippi Northern District	1959
Mississippi Southern District	1951
Missouri Eastern District	1932
Missouri Western District	1931
Montana District	1932
Nebraska District	1934
Nevada District	1933
New Hampshire District	1932
New Jersey District	1932
New Mexico District	1952
New York Eastern District	1930
New York Northern District	1939
New York Southern District	1932
New York Western District	1932
North Carolina Eastern District	1949
North Carolina Middle District	1949
North Carolina Western District	1949
North Dakota District	1948
Ohio Northern District	1943
Ohio Southern District	1945
Oklahoma Eastern District	1938

Oklahoma Northern District	1932
Oklahoma Western District	1930
Oregon District	1932
Pennsylvania Eastern District	1931
Pennsylvania Middle District	1945
Pennsylvania Western District	1931
Rhode Island District	1934
South Carolina District	1950
South Dakota District	1954
Tennessee Eastern District	1939
Tennessee Middle District	1945
Tennessee Western District	1956
Texas Eastern District	1951
Texas Northern District	1932
Texas Southern District	1932
Texas Western District	1931
Utah District	1950
Vermont District	1960
Virginia Eastern District	1934
Virginia Western District	1941
Washington Eastern District	1974
Washington Western District	1944
West Virginia Northern District	1934
West Virginia Southern District	1933
Wisconsin Eastern District	1950
Wisconsin Western District	1964
Wyoming District	1953
State Appellate Courts	
Alabama	1955
Alaska	1962
Arizona	1930
Arkansas	1950
California	1930
Colorado	1930
Connecticut	1950
Delaware	1950
Dist. of Columbia	1945
Florida	1910
Georgia	1940
Hawaii	1930
Idaho	1964

Illinois	1950
Indiana	1940
Iowa	1936
Kansas	1950
Kentucky	1943
Louisiana	1980
Maine	1940
Maryland	1949
Massachusetts	1930
Michigan	1930
Minnesota	1930
Mississippi	1954
Missouri	1960
Montana	1925
Nebraska	1965
Nevada	1865
New Hampshire	1930
New Jersey	1930
New Mexico	1930
New York	1848
North Carolina	1943
North Dakota	1930
Ohio	1950
Oklahoma	1954
Oregon	1950
Pennsylvania	1950
Rhode Island	1950
South Carolina	1950
South Dakota	1965
Tennessee	1950
Texas	1950
Utah	1950
Vermont	1930
Virginia	1930
Washington	1935
West Virginia	1970
Wisconsin	1945
Wyoming	1960 A143 A135 A145:B186