

# Legal Metadata Taxonomy Reference Guide

A Complete A–Z Reference for Law Firms, Legal Departments & Law Libraries

Published by Soutron Global | Updated 2026 | [soutron.com](https://soutron.com)

Legal metadata taxonomies are the structured classification frameworks that transform a law firm's document collection into a searchable, governable knowledge system. This guide provides a quick-reference table of established legal metadata taxonomies organised by functional category, followed by a full A–Z entry for each. Use it to audit your firm's current taxonomy coverage, identify gaps, and align internal schemas with recognised industry standards.

## QUICK REFERENCE: Legal Metadata Taxonomies by Function

Taxonomy	Primary Function	Key Use Case
SALI	Matter & service classification	Practice management, e-billing interoperability
Westlaw KeyNumber	Case law subject classification	Legal research, precedent retrieval
LCSH Law	Authoritative subject headings	Law library cataloging
BIALL	UK/Irish legal subject classification	Common law jurisdiction libraries
ECLI	EU case law identifier	Cross-border EU court record management
IPC / CPC	Patent subject classification	IP prosecution, FTO analysis
ISO 15489	Records management	Retention scheduling, regulatory compliance
NIEM	Government data exchange	Court systems, law enforcement interoperability
SKOS	Linked data / ontology	AI-ready taxonomy publishing
XBRL	Financial/regulatory disclosure	SEC filings, corporate governance
Dublin Core	General digital resource description	Open access legal repositories
Nice Classification	Trademark goods/services	Trademark prosecution, IP portfolio
PACER	Federal court document schema	Federal court filing metadata
MARC Relator	Legal role authority	Cataloging party and role metadata
UNODC	Criminal offense classification	Criminal law, international enforcement

## SUBJECT CLASSIFICATION & LEGAL RESEARCH TAXONOMIES

### **AALL Universal Citation Guide**

Standardizes citation formats across jurisdictions and document types. Functions as a taxonomy of legal document types and source categories used in authority control and cataloging in law libraries.

### **ABA Model Rules of Professional Conduct — Subject Classification**

The ABA's structured classification of professional responsibility topics, used as a controlled vocabulary in legal ethics research databases and law firm compliance libraries.

### **BIALL Classification Scheme**

Developed by the British and Irish Association of Law Libraries. Provides hierarchical subject classification for UK and Irish legal materials, with depth in English statute and case law. Particularly relevant for UK-headquartered firms and those operating across English, Scottish, Welsh, and Irish jurisdictions.

### **Black's Law Dictionary Controlled Vocabulary**

Functions as a de facto authority file for legal terminology. Widely used by legal information professionals as a source of preferred terms in subject indexing and thesaurus construction, ensuring consistency across teams and matter types.

### **Cornell LII Topic Taxonomy**

Cornell Law School's publicly available subject taxonomy organising U.S. law by topic, statute, regulation, and case category. Used in legal research portals and as a reference structure for subject heading alignment.

### **FindLaw Legal Topic Taxonomy**

A hierarchical subject classification used across Thomson Reuters consumer and professional legal research properties, organising legal topics by practice area, document type, and jurisdiction.

### **LCSH (Library of Congress Subject Headings) — Law**

One of the most comprehensive and widely applied legal taxonomies in existence, covering all areas of U.S. and international law with authorised headings, geographic subdivisions, and form subdivisions. Used in law firm libraries, academic law libraries, and court libraries worldwide.

### **LexisNexis Legal Taxonomy**

The proprietary subject classification schema underlying the LexisNexis research platform, organising case law, statutes, regulations, and secondary sources by practice area, jurisdiction, and legal topic.

### **LII Wex Legal Dictionary and Encyclopedia Taxonomy**

Cornell Law School's Wex reference database organised around a controlled subject taxonomy defining relationships between legal concepts, doctrines, and terms. Widely used as an authority reference for legal subject indexing.

### **Thomson Reuters Practical Law Taxonomy**

Subject classification schema used across Practical Law practice guides and standard documents, organising legal know-how by practice area, jurisdiction, and document type.

### **Westlaw Key Number System**

One of the oldest and most comprehensive legal taxonomies in existence, organising all reported U.S. case law into a hierarchical classification with numbered headnotes. In continuous development since the 1890s and a primary reference taxonomy for U.S. common law subject classification.

## **MATTER, CASE & COURT RECORD STANDARDS**

### **ECLI (European Case Law Identifier)**

A standardised identifier and metadata schema adopted by EU and member state courts to uniformly describe and cite court decisions across EU jurisdictions, covering court identity, jurisdiction, year, and case number in machine-readable format.

### **NIEM — Legal Domain (National Information Exchange Model)**

A U.S. federal framework providing a common vocabulary and data model for information exchange between government agencies, courts, and law enforcement. The legal domain defines standardised metadata fields for case records, party roles, charge classifications, and court event types.

### **PACER Metadata Schema**

The federal courts' PACER system uses a defined metadata schema for all filed documents, including case type, docket entry type, party roles, and filing dates. Functions as a de facto standard for U.S. federal court document metadata.

### **SALI (Standards Advancement for the Legal Industry)**

The most significant emerging legal metadata taxonomy designed specifically for the modern legal industry. Provides standardised matter-level codes covering practice areas, document types, industry sectors, legal entities, and geographic codes via the Legal Matter Standard Specification (LMSS) — a taxonomy of over 10,000 tags. Enables interoperability between practice management, e-billing, and document management systems.

## **INTELLECTUAL PROPERTY CLASSIFICATION FRAMEWORKS**

### **CPC / USPC (USPTO Classification System)**

The United States Patent and Trademark Office uses the Cooperative Patent Classification (CPC), a joint schema with the European Patent Office, as the controlling taxonomy for patent subject matter. Essential in patent prosecution, portfolio management, and freedom-to-operate analysis.

### **EUIPO Classification System**

Governs the classification of trademarks, designs, and IP rights across EU member states, organised by Nice Classification (goods and services), Vienna Classification (figurative elements), and Locarno Classification (industrial designs).

### **IPC (International Patent Classification)**

Administered by WIPO, the IPC provides a hierarchical classification of technology subject areas used globally to classify patent documents. Essential metadata in IP law practice and prosecution.

### **Nice Classification (NCL)**

The international classification of goods and services for trademark registration purposes, administered by WIPO. Mandatory metadata in trademark prosecution and IP portfolio management.

### **ULAN (Union List of Artist Names)**

Provides authority control for the identity of artists and cultural creators referenced in legal documents involving copyright, moral rights, and provenance. Used in arts law, IP, and cultural property practice.

### **WIPO Lex Legal Database Classification**

Classifies national and international IP legislation, treaties, and court decisions by type, subject matter, and jurisdiction. Used as a reference vocabulary in international IP practice metadata management.

## **RECORDS MANAGEMENT & ARCHIVAL METADATA STANDARDS**

### **Dublin Core Metadata Initiative — Legal Adaptation (DC-Law)**

An adaptation of the Dublin Core 15-element standard mapped to legal document description, used in legal digital libraries and open access repositories to ensure interoperability between systems.

### **EAD (Encoded Archival Description) — Legal Records**

An XML standard for encoding finding aids in archives. Law firm archives, court archives, and government legal record repositories use EAD with legal-specific controlled vocabulary extensions to describe collections of case files, corporate records, and regulatory submissions.

### **GARE (Guidelines for Authority Records and References)**

IFLA's guidelines for the construction of authority records, adapted in legal library contexts to manage name authority files for courts, legislators, law firms, and legal entities — ensuring consistent identification of legal actors across catalog records.

### **IFLA LRM (Library Reference Model) — Legal Collections**

Provides a conceptual framework for describing legal information resources, particularly useful in managing relationships between legal texts, their editions, translations, and commentary.

### **ISO 5127 — Information and Documentation Vocabulary**

An ISO standard providing a controlled vocabulary for information and documentation concepts, including legal records management terminology. Used as a reference framework in legal information governance and records management policy.

### **ISO 15489 — Records Management Metadata Standard**

The foundational international standard for records management, governing the creation, capture, classification, retention, and disposal of records. Widely applied in corporate legal departments and law firm records management programs operating under GDPR, CCPA, or SEC recordkeeping rules.

### **ISO 19475 — Legal Document Exchange Standard**

Provides structured metadata requirements for the electronic exchange of legal documents between parties, courts, and regulatory bodies, with defined fields for document type, party identification, and procedural status.

### **MARC Relator Codes — Legal Roles**

The MARC standard includes relator codes defining the roles of persons and corporate bodies in relation to a document. In legal cataloging, these describe roles such as attorney of record, court, plaintiff, defendant, intervenor, and amicus.

### **OCLC Dewey Decimal Classification — Law (340s)**

The Law schedule (340–349) provides a hierarchical numerical taxonomy for organising legal collections by jurisdiction and subject, used in law firm libraries, court libraries, and public law collections.

## **COMPLIANCE, REGULATORY & FORENSIC TAXONOMIES**

### **ICD — Forensic and Medico-Legal Application**

In forensic, personal injury, and workers' compensation legal practice, ICD diagnostic codes function as a controlled vocabulary for classifying injury types, causes of death, and medical conditions referenced in legal documents and damage calculations.

### **MeSH — Forensic and Legal Medicine Application**

The National Library of Medicine's Medical Subject Headings are used in legal medicine, forensic pathology, and personal injury practice to provide controlled vocabulary for medical subject matter referenced in legal documents.

### **UNODC Legal Classification**

Provides a standardised taxonomy for criminal offense classification, drug scheduling categories, and international crime typologies. Used in criminal law practice, international enforcement cooperation, and forensic records management.

### **XBRL — Legal and Regulatory Filing Taxonomy**

Provides a standardised taxonomy for structured financial and regulatory disclosure data, used in SEC filings, corporate governance documents, and legal compliance reporting. XBRL taxonomy codes function as metadata fields linking legal documents to specific regulatory disclosure requirements.

## **INTEROPERABILITY, LINKED DATA & AI-READY STANDARDS**

### **CILIP Thesaurus for Graphical Materials — Legal Adaptation**

Used in archival and special collections contexts where legal records intersect with visual or documentary materials. Particularly relevant for legal archives managing photographic evidence, architectural records, maps, and visual exhibits.

### **SKOS (Simple Knowledge Organization System) — Legal Ontology**

A W3C standard for representing controlled vocabularies, thesauri, and classification schemes in machine-readable format. Used in legal knowledge graph and ontology projects — including the EU Publications Office's EuroVoc and the LKIF (Legal Knowledge Interchange Format) — to publish and interlink legal taxonomies as linked data. Key standard for AI-ready taxonomy publishing.

### **UN/LOCODE and UN Trade Terms — Trade and Customs Law**

United Nations location codes and trade term classifications (including Incoterms) function as controlled vocabularies in international trade law, customs compliance, and cross-border transaction metadata.

## Ready to Build a Smarter Legal Metadata Strategy?

Soutron's purpose-built legal library platform supports all of these taxonomy frameworks within a single, unified polyhierarchical thesaurus — so your documents are classified everywhere they should be and discoverable from every relevant search path. No single-parent hierarchy limitations. No retrieval failures caused by taxonomic ambiguity.

Talk to our team about aligning your firm's metadata architecture with the standards that matter most to your practice.

**[Book a free consultation at soutron.com/demo](https://soutron.com/demo)**