

Legal Taxonomy and Metadata Reference Guide

A verified A to Z reference for law firms, legal departments and law libraries

Published by Soutron Global | October 2026 | soutron.com

A legal taxonomy is the controlled vocabulary a legal organization uses to classify its own documents, matters and records. This guide describes the published standards, schemes and models most often cited in that work, with a routing table that maps the job you are doing onto the entries worth examining first.

How this guide is compiled. Every entry is checked against the issuing body's own documentation for its official name, publisher, current status and what it actually governs. Only resources that exist as named are listed. Where a resource is widely described as a legal taxonomy but is in fact something else, a database, a dictionary or a set of rules, this guide says so. A section near the end covers the errors that recur most often in published summaries of this subject.

Following a standard to its source. Every entry name in the A to Z is a link to the issuing body's own page for that standard. The full addresses are listed together in a table at the end for anyone reading a printed copy.

QUICK REFERENCE: Which Standard Does Which Job

The job you are doing	Standard to examine	Note
Classifying matters and legal work by type	SALI Legal Matter Specification Standard	Shared terminology so a firm and its client describe the same work the same way
Classifying a law library collection	LC Classification Class K, Moys, Dewey 340s	Class K and Moys are the two schemes most law libraries choose between
Applying subject headings to legal material	Library of Congress Subject Headings	A subject vocabulary, not a classification, and a separate product from Class K
Indexing case law by legal issue	West Key Number System, LexisNexis Legal Taxonomy	Proprietary and platform-tied. Mapping targets rather than schemes to adopt
Citing decisions consistently	ECLI, AALL Universal Citation Guide	Citation standards rather than subject schemes
Exchanging data with courts and justice agencies	NIEMOpen Justice domain, OASIS LegalXML ECF	NIEMOpen supplies the vocabulary. ECF is the filing standard built on it
Describing archival and historic firm material	ISAD(G), EAD, Dublin Core	Rules for description, XML encoding of the finding aid, and a general element set
Classifying intellectual property subject matter	IPC, CPC, Nice, the classifications EUIPO applies	Each covers a different right
Researching IP legislation and decisions	WIPO Lex	A searchable database, not a classification system
Publishing a vocabulary for other systems	SKOS	A data model for exchange, not a subject vocabulary
Records retention and disposition	ISO 15489-1, with ISO 23081 for metadata	Governs the lifecycle, not the subject terms
Regulatory and financial disclosure filings	XBRL	Uses taxonomies in a technical sense. Securities practice rather than the library

Two cautions. Adopting a published standard does not remove the need for local decisions, because every entry still has to be mapped onto your own material. And a scheme that fits one part of a collection rarely fits all of it, so the realistic goal is a small number of schemes that interoperate rather than one that wins.

LAW LIBRARY CLASSIFICATION AND SUBJECT ACCESS

Library of Congress Classification, Class K (Law)

The hierarchical shelf classification for legal materials, running across subclasses K to KZ and covering law in general, comparative law, jurisprudence, religious law and the laws of individual jurisdictions. It is the scheme most academic and large law libraries use, and it is a different product from Library of Congress Subject Headings.

Library of Congress Subject Headings (LCSH)

A single alphabetical controlled vocabulary maintained by the Library of Congress Policy and Standards Division and in use since 1898. It has no distinct law section: law headings are interspersed alphabetically with all others. LCSH describes what an item is about, while Class K decides where it sits on the shelf, and the two are routinely confused.

Moys Classification and Thesaurus for Legal Materials

A hierarchical classification scheme and thesaurus for legal collections, created by Elizabeth M. Moys in 1968 and designed to sit alongside Library of Congress Classification. The current fifth edition, revised by Diana Morris, was published by De Gruyter Saur in 2012 under the auspices of the British and Irish Association of Law Librarians, and the scheme is used in law libraries across common law jurisdictions including the UK, Ireland, Canada, Australia and New Zealand.

Dewey Decimal Classification, Law (340 to 349)

The Dewey Decimal Classification, owned and maintained by OCLC, assigns law to the 340s, running from 340 Law and 341 Law of nations through 345 Criminal law and 347 Civil procedure and courts to 349 Law of specific jurisdictions. It is a general library classification with a law range rather than a legal scheme, which is why most dedicated law libraries use Moys or Class K instead.

MARC Code List for Relators

Maintained by the Library of Congress Network Development and MARC Standards Office, this is a controlled list of three-character codes and terms describing the role a person or corporate body plays in relation to a resource. It includes legal roles such as defendant, plaintiff, appellant, appellee, judge and witness, listed alphabetically among all other relators rather than grouped as a named legal subset.

Thesaurus for Graphic Materials (TGM)

Compiled and maintained by the Prints and Photographs Division of the Library of Congress, TGM holds roughly 7,500 authorised terms for indexing the subjects depicted in pictorial material. It is included here because it is commonly miscited in legal information contexts: it has no connection to CILIP and no legal adaptation, and it is relevant only where a legal collection includes photographic or graphic holdings.

MATTER, CASE LAW AND CITATION STANDARDS

SALI Alliance and the Legal Matter Specification Standard (LMSS)

The SALI Alliance (Standards Advancement for the Legal Industry) is a not-for-profit standards body founded in 2017 by the Association of Legal Administrators, the International Legal Technology Association and the Legal Marketing Association. It publishes the Legal Matter Specification Standard, an RDF and OWL based taxonomy for classifying legal matters, areas of law, industries and related attributes. Version 2.0 was released in 2022, a SALI API standard followed in 2024, and version 3 is being published attribute by attribute.

West Key Number System

Owned by Thomson Reuters and originating with West Publishing in the nineteenth century, the Key Number System classifies points of law into more than 400 topics and roughly 100,000 key numbers, applied to editorial headnotes on cases in Westlaw and the West digest and reporter system. It is a publisher's editorial product rather than an open standard, so it is a mapping target rather than a scheme to adopt.

LexisNexis Legal Taxonomy

The proprietary subject classification underlying the LexisNexis research platform, surfaced through Lexis headnotes and the legal topic hierarchy used in Search by Topic and Search by Headnote. It is not published as a browsable public vocabulary, so treat it as a mapping target rather than a scheme to adopt.

AALL Universal Citation Guide

A set of vendor-neutral citation rules for judicial decisions, statutes and administrative regulations, recommended by the American Association of Law Libraries and published in its third edition (2014) by William S. Hein and Co. It is a citation standard rather than a subject classification, so it governs how a decision is cited, not how it is indexed.

ECLI (European Case Law Identifier)

A standardised identifier and an accompanying minimum set of uniform metadata for case law, based on Dublin Core, introduced by Council of the European Union conclusions in 2011 and revised in 2019. Adoption is voluntary by member state, each appoints a national ECLI coordinator, and the European Commission runs the ECLI search engine on the e-Justice Portal.

Cornell Legal Information Institute and Wex

The Legal Information Institute at Cornell Law School publishes Wex, a free legal dictionary and encyclopedia with a browsable subject index of roughly 200 legal topics. It is a navigational subject listing rather than a published controlled vocabulary, and it is not distributed as a structured taxonomy, so treat it as a reference source rather than a scheme to adopt.

DATA EXCHANGE AND INTEROPERABILITY STANDARDS

NIEMOpen and its Justice domain

Formerly the National Information Exchange Model and now an OASIS Open Project, NIEMOpen provides a common vocabulary and reference data model for information exchange across government agencies. Its Justice domain carries the components for criminal justice, court case management and legal process data, and the current release is NIEM Model version 6.0, approved as an OASIS Standard in November 2025. There is no domain called Legal.

OASIS LegalXML Electronic Court Filing (ECF)

The operative standard for electronic court filing interoperability, defining the XML and web services specifications used to exchange filings between filing applications and court systems. Version 5.0 was approved in April 2019 and version 5.01 in September 2023. ECF builds on NIEM, which is why the two are often mentioned together.

SKOS (Simple Knowledge Organization System)

A W3C Recommendation published on 18 August 2009, SKOS provides an RDF vocabulary and data model for representing controlled vocabularies, thesauri, taxonomies and classification schemes in machine-readable form. It is used by legal and legislative vocabularies including EuroVoc from the Publications Office of the European Union, though no SKOS legal profile exists as a published standard. When a vendor says a system supports SKOS, that is a statement about exchange, not about the quality of the vocabulary inside it.

INTELLECTUAL PROPERTY CLASSIFICATION

IPC (International Patent Classification)

Administered by WIPO under the Strasbourg Agreement of 1971, the IPC is a hierarchical, language-independent system of sections, classes, subclasses and groups used worldwide to classify patents and utility models. A new version takes effect each 1 January.

Cooperative Patent Classification (CPC)

Developed and maintained jointly by the United States Patent and Trademark Office and the European Patent Office, CPC is the controlling classification for patent subject matter and is used by patent offices worldwide. The older United States Patent Classification is static for utility patents, which stopped receiving USPC codes after 31 December 2014, and is now updated only for design and plant classes.

Nice Classification (NCL)

The international classification of goods and services for trademark registration, established by the Nice Agreement of 1957 and administered by WIPO. A new edition takes effect each 1 January.

EUIPO and the classifications it applies

The European Union Intellectual Property Office does not publish a classification of its own. It applies three WIPO-administered systems, Nice for goods and services, Vienna for figurative trade mark elements, and Locarno for industrial designs, and co-maintains with EU national offices the Harmonised Database of pre-approved goods and services terms, accessible through TMclass, DesignClass and EuroLocarno.

WIPO Lex

A free database maintained by the World Intellectual Property Organization, giving access to more than 50,000 legal documents from around 200 jurisdictions across three collections: IP laws and regulations, treaties, and judicial decisions. It offers browse and filter facets by subject matter and jurisdiction, but WIPO publishes no named classification scheme for it. WIPO's named classifications are IPC, Nice, Vienna and Locarno.

RECORDS MANAGEMENT AND ARCHIVAL DESCRIPTION

ISO 15489-1 and ISO 23081

ISO 15489-1:2016 sets out the concepts and principles for creating, capturing and managing records, including policy, responsibilities, records controls and records processes. It is not a metadata standard. The metadata requirements sit in ISO 23081, currently parts 1:2017 (Principles), 2:2021 (Conceptual and implementation issues) and 3:2011 (Self-assessment method).

ISO 5127, Information and documentation, Foundation and vocabulary

A terminology standard that defines a concept system and general vocabulary for documentation, libraries, archives, media, museums, records management and conservation. Clause 3.13 covers legal aspects of information and documentation, chiefly intellectual property, author's rights, privacy and data protection, rather than legal records terminology.

EAD (Encoded Archival Description)

An XML standard for encoding archival finding aids, maintained by the Society of American Archivists' Technical Subcommittee on Encoded Archival Standards, with the Library of Congress hosting the official schemas. The current version is EAD 4, and legal repositories use it to describe court records, government legal archives and law firm collections, including the legal status of records.

Dublin Core

Maintained by the Dublin Core Metadata Initiative, Dublin Core provides a general-purpose set of 15 core metadata elements, now part of DCMI Metadata Terms, widely used for describing documents in repositories including legal collections. DCMI publishes no legal profile, although individual projects have mapped Dublin Core to legal material, including Cornell LII's documented mapping for judicial opinion metadata.

IFLA LRM (Library Reference Model)

Published by IFLA in 2017, LRM is a high-level conceptual model for bibliographic information that consolidates FRBR, FRAD and FRSAD into a single model. It is a conceptual framework rather than a metadata schema or vocabulary, and it reaches legal collections indirectly, through RDA, the cataloguing standard built on it.

ADJACENT CLASSIFICATIONS USED IN LEGAL PRACTICE

ICD (International Classification of Diseases)

The International Statistical Classification of Diseases and Related Health Problems is maintained by the World Health Organization, with ICD-11 in force since January 2022. It is not a legal standard, but its codes function as a controlled vocabulary for injury types and causes of death recorded on death certificates and referenced in personal injury, workers' compensation and coronial matters.

ICCS (International Classification of Crime for Statistical Purposes)

Published by the United Nations Office on Drugs and Crime and endorsed in 2015 by the UN Statistical Commission and the Commission on Crime Prevention and Criminal Justice, ICCS version 1.0 classifies criminal offences by event type so that crime and criminal justice data can be compared across countries. It is a statistical classification rather than a legal information scheme, and it does not cover drug scheduling.

XBRL (eXtensible Business Reporting Language)

An open reporting standard governed by XBRL International, in which taxonomy is a term of art meaning a machine-readable dictionary of reporting concepts rather than a subject classification. For US securities filings the relevant taxonomies are the US GAAP Financial Reporting Taxonomy and the SEC Reporting Taxonomy, maintained by the Financial Accounting Foundation and required by the SEC for EDGAR filings.

COMMONLY MISQUITED IN SUMMARIES OF THIS SUBJECT

These errors recur across published lists of legal metadata standards and software vendor pages. Each entry below was checked against the issuing body.

Frequently stated	Actually
Incoterms are United Nations trade terms	Incoterms are published by the International Chamber of Commerce, currently Incoterms 2020, first published 1936. UNCITRAL recognises them but does not publish them. UN/LOCODE is a separate UNECE transport location code list
ISO 19475 is a legal document exchange standard	ISO 19475:2021 covers document management and the minimum requirements for the storage of documents. It has no legal exchange scope. Its three 2018 parts are withdrawn
PACER publishes a metadata schema	PACER is a public access service run by the Administrative Office of the US Courts. The court filing data standard is OASIS LegalXML Electronic Court Filing
ISO 15489 is the records metadata standard	ISO 15489-1 covers records management concepts and principles. The metadata standard is ISO 23081
LCSH has a law section	It does not. Law headings are interspersed alphabetically. The hierarchical legal scheme is Library of Congress Classification Class K
The BIALL classification scheme	The scheme is the Moys Classification and Thesaurus for Legal Materials. BIALL sponsors it, and is neither its author nor its publisher
NIEM has a Legal domain	NIEM has a Justice domain. There is no domain called Legal
The CILIP Thesaurus for Graphical Materials	The Thesaurus for Graphic Materials, published by the Library of Congress. CILIP has no involvement
Black's Law Dictionary is a controlled vocabulary	It is a dictionary of legal definitions. No authority file or indexing vocabulary is derived from it
Dewey is no longer owned by OCLC	OCLC owns and maintains the Dewey Decimal Classification

WHERE TO FIND THESE STANDARDS

Each address below belongs to the body that issues or maintains the standard. Checked October 2026.

Standard	Issuing body page
Library of Congress Classification, Class K (Law)	www.loc.gov/aba/publications/FreeLCC/freelcc.html
Library of Congress Subject Headings (LCSH)	www.loc.gov/aba/publications/FreeLCSH/freelcsh.html
Moys Classification and Thesaurus for Legal Materials	biall.org.uk/publications-2
Dewey Decimal Classification, Law (340 to 349)	www.oclc.org/en/dewey.html
MARC Code List for Relators	www.loc.gov/marc/relators
Thesaurus for Graphic Materials (TGM)	guides.loc.gov/tgm-i
SALI Alliance and the Legal Matter Specification Standard (LMSS)	www.sali.org
West Key Number System	www.thomsonreuters.com/en-us/help/westlaw-classic/tools/access-the-west-key-number-system
LexisNexis Legal Taxonomy	supportcenter.lexisnexis.com/app/answers/answer_view/a_id/1094669
AALL Universal Citation Guide	www.aallnet.org/resources-publications/publications/universal-citation-guide
ECLI (European Case Law Identifier)	e-justice.europa.eu/topics/legislation-and-case-law/european-case-law-identifier-ecli_en
Cornell Legal Information Institute and Wex	www.law.cornell.edu/wex
NIEMOpen and its Justice domain	niemopen.org
OASIS LegalXML Electronic Court Filing (ECF)	docs.oasis-open.org/legalxml-courtfilling/ecf/v5.01/ecf-v5.01.html
SKOS (Simple Knowledge Organization System)	www.w3.org/TR/skos-reference
IPC (International Patent Classification)	www.wipo.int/en/web/classification-ipc
Cooperative Patent Classification (CPC)	www.uspto.gov/patents/search/classification-standards-and-development
Nice Classification (NCL)	www.wipo.int/en/web/classification-nice
EUIPO and the classifications it applies	www.euipo.europa.eu/en/help-centre/searches/faq-harmonised-database
WIPO Lex	www.wipo.int/en/web/wipolex
ISO 15489-1 and ISO 23081	www.iso.org/standard/62542.html
ISO 5127, Information and documentation, Foundation and vocabulary	www.iso.org/standard/59743.html
EAD (Encoded Archival Description)	www.loc.gov/ead
Dublin Core	www.dublincore.org/specifications/dublin-core
IFLA LRM (Library Reference Model)	www.ifla.org/ifla-s-bibliographic-conceptual-models
ICD (International Classification of Diseases)	www.who.int/standards/classifications/classification-of-diseases
ICCS (International Classification of Crime for Statistical Purposes)	unstats.un.org/unsd/classifications/Family/Detail/1000
XBRL (eXtensible Business Reporting Language)	www.xbrl.org/the-standard/what/key-concepts-in-xbrl/taxonomies

FOUR QUESTIONS TO ASK BEFORE YOU COMMIT TO A SCHEME

Question	What the answer tells you
Can one concept sit under more than one branch of the hierarchy?	Whether you have a thesaurus or a single tree. Legal material routinely belongs in several places at once, and a strict tree forces a choice that loses one of them
Who can add or rename a term, and does it require a developer?	Whether the vocabulary is configuration or code. A taxonomy nobody can edit without a release cycle stops being maintained within a year
What happens to records already indexed when a term is renamed or merged?	Whether term changes carry through to existing records or leave them stranded under a label that no longer exists
Can the vocabulary be exported in a standard form such as SKOS?	Whether the work you put into the scheme is portable, or tied to the system that holds it

Working out how your classification should be structured?

Choosing standards is one problem. Making them hold together across practice areas, offices and the systems you already run is a different one. Soutron's legal library software treats classification as something a firm configures rather than something it inherits from the software. Controlled vocabularies are maintained through validation lists and thesaurus management, and a poly-hierarchical thesaurus allows a single concept to sit under more than one branch of a taxonomy, so a memo that belongs in several places is findable from each of them. Firms working to industry terminology can align their taxonomies with established legal standards, including frameworks from the SALI Alliance.

Book a consultation

soutron.com/demo