

هستی‌شناسی حقوقی و بازیابی معنایی اطلاعات: مرور نظام‌مند رویکردها و چارچوب بومی برای ایران

سهیلا نبی زاده کیوی

دانشجوی دکتری علم اطلاعات و دانش‌شناسی (مدیریت اطلاعات و دانش)؛ دانشکده علوم تربیتی و روانشناسی؛ دانشگاه تبریز؛
تبریز؛ ایران. nabizadeh.64@gmail.com

رسول زوارقی*

دکتری علم اطلاعات و دانش‌شناسی؛ استاد؛ دانشکده علوم تربیتی و روانشناسی؛ دانشگاه تبریز؛ تبریز؛ ایران. پدیدآور رابط
zavaraqi@tabrizu.ac.ir

هاشم عطاپور

دکتری علم اطلاعات و دانش‌شناسی؛ دانشیار؛ دانشکده علوم تربیتی و روانشناسی؛ دانشگاه تبریز؛ تبریز؛ ایران.
hashematapour@tabrizu.ac.ir

آیت مولائی

دکتری حقوق؛ دانشیار؛ دانشکده حقوق و علوم اجتماعی؛ دانشگاه تبریز؛ تبریز؛ ایران. amulaee@tabrizu.ac.ir

مقدمه: رشد شتابان داده‌های حقوقی و ناهمگنی مفهومی قوانین و آرای قضایی، نیاز به روش‌های پیشرفته برای بازیابی معنایی اطلاعات حقوقی را برجسته می‌کند. جست‌وجوی واژگانی در مواجهه با هم‌معنایی، چندمعنایی و پیچیدگی ساختار حقوقی کارآمد نیست و نبود هستی‌شناسی‌های بومی در ایران این چالش را تشدید می‌کند. در سطح جهانی، رویکردهای مبتنی بر هستی‌شناسی حقوقی و استانداردهای وب معنایی نقش مهمی در ارتقای دقت بازیابی یافته‌اند. این پژوهش با بهره‌گیری از چارچوب زمینه-سازوکار-پیامد (CMO)، چهار پرسش اصلی را دنبال می‌کند: عوامل زمینه‌ای مؤثر بر طراحی هستی‌شناسی‌های حقوقی؛ سازوکارهای ارتقای بازیابی معنایی؛ الگوهای غالب و چالش‌های طراحی در نظام‌های حقوق موضوعه؛ و چارچوب مفهومی مناسب برای توسعه مدل‌های بومی در ایران. هدف پژوهش ارائه چارچوبی مفهومی برای بومی‌سازی هستی‌شناسی حقوقی در ایران است.

روش‌شناسی: این پژوهش با رویکرد مرور نظام‌مند و سنتز تفسیری انجام شد تا الگوهای «زمینه، سازوکار و پیامد» در طراحی و کاربرست هستی‌شناسی‌های حقوقی شناسایی گردد. فرایند مرور بر اساس دستورالعمل PRISMA 2020 ساختار یافت و شامل تدوین پرسش‌ها، طراحی پروتکل، جست‌وجوی منابع، غربالگری، ارزیابی کیفیت، استخراج داده‌ها و سنتز نهایی بود. جست‌وجوی نظام‌مند در پایگاه‌های Web of Science، Scopus، ACM Digital Library، Google Scholar و پایگاه‌های داخلی انجام شد و ۹۲۳ رکورد شناسایی گردید. پس از حذف موارد تکراری، ۶۸۲ رکورد برای غربالگری باقی ماند و در نهایت ۲۸ مطالعه بر اساس معیارهای ورود انتخاب شد. ارزیابی کیفیت با ابزار MMAT انجام شد و ضریب پایایی $\text{Cohen's Kappa} = 0.87$ به دست آمد.

داده‌ها در پنج مرحله شامل کدگذاری اولیه، استخراج تم‌ها، تحلیل بافت حقوقی، شناسایی سازوکارها و استخراج پیامدها سنتز شدند و مبنای تدوین چارچوب مفهومی بومی سازی شده قرار گرفتند.

یافته های اصلی: یافته‌های پژوهش که حاصل سنتز تفسیری ۲۸ مطالعه بین سال‌های ۱۹۹۷ تا ۲۰۲۵ بود، نشان داد که هستی‌شناسی‌های حقوقی طی سه دهه گذشته مسیر تحولی چشمگیر را طی کرده‌اند. در پاسخ به پرسش نخست، تحلیل‌ها آشکار ساخت که نوع نظام حقوقی، ویژگی‌های زبانی-فرهنگی و حوزه تخصصی حقوقی سه عامل اصلی تعیین‌کننده در اولویت‌های طراحی هستند. در پاسخ به پرسش دوم، پنج سازوکار کلیدی شناسایی شد: انتخاب هستی‌شناسی بالایی مناسب، گزینش روش ساخت (بالا به پایین، پایین به بالا، ترکیبی، میان به بالا)، ترکیب با فناوری‌های نوین مانند گراف‌های دانش و مدل‌های زبانی بزرگ، استفاده از نسخه‌های سبک برای کاربردهای مقیاس‌پذیر، و بهره‌گیری از الگوریتم‌های خودکار برای کاهش هزینه و زمان. در پاسخ به پرسش سوم، چهار الگوی منطقه‌ای (اروپایی، آسیای شرقی، آمریکای لاتین و کشورهای اسلامی-عربی) شناسایی شد که هر یک مسیر متفاوتی را در توسعه هستی‌شناسی پیموده‌اند. با وجود این تفاوت‌ها، چالش‌های مشترکی همچون بافت‌باز بودن مفاهیم حقوقی، شکاف میان نظریه حقوقی و مهندسی هستی‌شناسی، پویایی قوانین و هزینه بالای ساخت دستی مشاهده شد. در پاسخ به پرسش چهارم، نتایج نشان داد که چارچوب پیشنهادی برای ایران باید بر LKIF بومی‌سازی شده، بهره‌گیری از تجارب کشورهای مشابه، رویکرد ترکیبی مبتنی بر قانون مدنی و طراحی هستی‌شناسی چرخشی استوار باشد. این چارچوب می‌تواند گذار از جست‌وجوی کلیدواژه‌ای به بازیابی معنایی دقیق را در نظام حقوقی ایران امکان‌پذیر سازد.

بحث و نتیجه گیری: این مرور نظام‌مند نشان می‌دهد که هستی‌شناسی‌های حقوقی ساختارهایی بافت‌محورند و موفقیت آن‌ها وابسته به سازگاری با زمینه‌های حقوقی، فرهنگی و زبانی است. چارچوب «زمینه-سازوکار-پیامد» که در این پژوهش به کار رفته، امکان تحلیل تطبیقی رویکردهای جهانی را فراهم می‌کند و به عنوان الگویی روش‌شناختی برای مطالعات آینده ارزش دارد. یافته‌ها آشکار می‌سازد که روندهای نوظهوری همچون رویکردهای ترکیبی، مدل‌سازی تغییرات قانونی، معماری‌های سبک و ادغام با فناوری‌های نوین—از جمله مدل‌های زبانی بزرگ، داده‌های پیوندی و گراف‌های دانش—نقش محوری در ارتقای دقت، مقیاس‌پذیری و کاربرپذیری هستی‌شناسی‌ها دارند. تحلیل تطبیقی نشان می‌دهد که نظام‌های حقوق موضوعه نیازمند تمرکز بر ساختار هنجاری قوانین‌اند و چالش‌هایی مانند بافت‌باز بودن مفاهیم حقوقی، شکاف میان نظریه حقوقی و مهندسی هستی‌شناسی، پویایی قوانین و هزینه بالای ساخت دستی، ضرورت طراحی بومی را برای ایران برجسته می‌کند. این پژوهش استنباط می‌کند که بهره‌گیری از LKIF بومی‌سازی شده، رویکردهای ترکیبی مبتنی بر قانون مدنی و مدل‌سازی چرخشی قوانین و آرای وحدت رویه می‌تواند مسیر توسعه سامانه‌های بازیابی معنایی دقیق و عادلانه را فراهم سازد. پژوهش‌های آینده می‌توانند بر طراحی هستی‌شناسی قانون مدنی ایران، توسعه گراف دانش حقوقی فارسی، ارزیابی مدل‌های زبانی بزرگ در حقوق فارسی، و ایجاد سامانه‌های پرسش و پاسخ و بازیابی معنایی مبتنی بر هستی‌شناسی متمرکز شوند.

کلیدواژه‌ها: هستی‌شناسی حقوقی، بازیابی معنایی اطلاعات، مرور نظام‌مند، نظام حقوق موضوعه، تحلیل زمینه-سازوکار-پیامد (CMO)

Extended Abstract

Legal Ontologies for Semantic Information Retrieval: A Systematic Review and a Conceptual Framework for Civil Law Systems

Soheila Nabizadeh Kivi

PhD Candidate in Knowledge and Information Science (Knowledge and Information Management),
Faculty of Educational Sciences and Psychology, University of Tabriz, Tabriz, Iran. Email:
nabizadeh.64@gmail.com

Rasoul Zavaraqi*

PhD in Knowledge and Information Science; Professor, Faculty of Educational Sciences and
Psychology, University of Tabriz, Tabriz, Iran. Corresponding Author. Email: zavaraqi2024@gmail.com

Hashem Atapour

PhD in Knowledge and Information Science; Associate Professor, Faculty of Educational Sciences
and Psychology, University of Tabriz, Tabriz, Iran. Email: hashematapour@tabrizu.ac.ir

Ayat Mulaee

PhD in Law; Associate Professor, Faculty of Law and Social Sciences, University of Tabriz, Tabriz,
Iran. Email: amulaee@tabrizu.ac.ir

Introduction: The rapid growth of legal data and the conceptual heterogeneity of statutes and judicial decisions highlight the need for advanced methods of semantic legal information retrieval. Keyword-based search is ineffective when facing synonymy, polysemy, and the structural complexity of legal texts, and the absence of indigenous legal ontologies in Iran intensifies this challenge. Globally, ontology-based approaches and Semantic Web standards have significantly improved retrieval accuracy. Using the Context–Mechanism–Outcome (CMO) framework, this study addresses four key questions: the contextual factors influencing the design of legal ontologies; the mechanisms that enhance semantic retrieval; the dominant patterns and challenges in civil-law systems; and the conceptual framework required for developing indigenous models in Iran. The aim is to propose a conceptual foundation for localizing legal ontologies in the Iranian legal system.

Methodology: This study employed a systematic review and interpretive synthesis to identify patterns of context, mechanism, and outcome in the design and application of legal ontologies. The review followed PRISMA 2020 guidelines and included formulating research questions, designing the protocol, searching databases, screening, quality appraisal, data extraction, and final synthesis. A comprehensive search across Web of Science, Scopus, ACM Digital Library, Google Scholar, and national databases retrieved 923 records, of which 28 studies met the inclusion criteria. Quality appraisal used the Mixed Methods Appraisal Tool (MMAT 2018), and inter-rater reliability was confirmed (Cohen's Kappa = 0.87). Data were synthesized through five steps: initial coding, theme extraction, contextual analysis, mechanism identification, and outcome mapping, forming the basis for a localized conceptual framework.

Main Findings: The synthesis of 28 studies (1997–2025) shows that legal ontologies have

undergone substantial evolution over the past three decades. Regarding the first question, legal system type, linguistic-cultural features, and legal domain specificity emerged as the main contextual determinants of ontology design. For the second question, five core mechanisms were identified: selecting an appropriate upper ontology, choosing a construction method (top-down, bottom-up, hybrid, middle-out), integrating modern technologies such as knowledge graphs and large language models, using lightweight architectures, and employing automated algorithms to reduce cost and time. For the third question, four regional patterns—European, East Asian, Latin American, and Islamic-Arab systems—were identified, each following distinct developmental trajectories. Despite these differences, common challenges persist: open-textured legal concepts, the theory-engineering gap, legal dynamism, and the high cost of manual ontology construction. For the fourth question, the proposed framework for Iran relies on localized LKIF, lessons from comparable jurisdictions, a hybrid approach grounded in the Civil Code, and rotational modeling of statutes and judicial precedents.

Discussion and Conclusions: This review demonstrates that legal ontologies are inherently context-dependent, and their success relies on alignment with legal, cultural, and linguistic environments. The CMO framework provides a robust methodological lens for comparative analysis and future research. Emerging trends—including hybrid approaches, modeling legal change, lightweight architectures, and integration with modern AI technologies such as LLMs, linked data, and knowledge graphs—play a central role in enhancing accuracy, scalability, and applicability. Civil-law systems require ontology designs focused on normative structures, and challenges such as open-textured concepts, the theory-engineering gap, and legal dynamism underscore the need for indigenous development in Iran. The study concludes that localized LKIF, hybrid modeling based on the Civil Code, and rotational modeling of statutes and judicial precedents can pave the way for precise and equitable semantic retrieval. Future research may focus on developing an ontology for the Iranian Civil Code, constructing a Persian legal knowledge graph, evaluating LLMs for Persian legal tasks, and designing ontology-based semantic retrieval and question-answering systems.

Keywords: Legal ontology, Semantic information retrieval, Systematic review, Civil-law systems, Context-Mechanism-Outcome (CMO) analysis