

法律资源的语义模型构建研究

邢启迪, 耿骞, 赵盼云, 景然
(北京师范大学政府管理学院, 北京 100875)

摘要: 本文首先从法理学视角, 界定我国法律领域各类法律资源类型; 其次, 从信息资源组织角度, 对各类法律资源进行深入分析, 识别出可描述的实体属性和实体关系; 再次, 使用frbr、metalex、dct等本体和本文构建的linkedlaw法律本体, 对各类法律资源设计描述框架; 最后, 依据各类法律资源间的语义关联关系, 在各类法律资源间建立链接, 进而构建一个法律资源语义模型。本文实例表明, 该语义模型能够全面组织和关联各类法律资源, 生成一张多层次、多粒度、多类型、全方位的关联法律数据网, 更好地满足法律信息需求。

关键词: 法律信息组织; 关联法律数据; 法律本体

中图分类号: G354.4

DOI: 10.3772/j.issn.1673-2286.2017.05.006

1 引言

法律数据是政府开放数据的重要组成部分。随着政府信息化和政府数据开放工作的开展和深入, 越来越多的法律数据可通过Web访问获取。如中央政府网站发布的法律法规, 中国裁判文书网发布的各级法院裁判文书, 国务院各部委以及各级地方政府网站发布的相关法规和规章, 一些商业法律数据库提供的法律法规和裁判文书检索与传递服务, 工商总局发布的注册企业基本信息等。但目前用户在获取和利用法律信息过程中, 仍会遇到三方面问题。

第一, 信息组织仍然以文档或文件为单位, 而未以数据为单位, 组织和标引方式简单, 信息粒度粗放。以法律文献资源为例, 多数网站以整部法律为单位进行组织和展示, 用户很难直接定位局部信息。另外, 法律法规由多个法律条款组成, 裁判文书会引用法律条款, 法律条款间也会相互引用, 用户经常要在不同法律条款和裁判文书文本间按照特定路径获取相关信息。但目前裁判文书在引用法律条款时, 通常仅提及法律的名称和被引用条款在该法律中的序号, 并不插入该条款的具体文字内容。此外, 一部法律法规通常有多个版本, 而多数情况下裁判文书不提及所引用法律法规的版本信息, 用户只能依据裁判文书的发布时间和当时正在实施的法律版本进行人工判断。

第二, 各类法律信息资源间缺少语义关联关系。例如, 目前政府相关部门以及专业法律数据提供商都提供法律法规、裁判文书等, 但这类资源相对孤立, 缺少与诉讼当事人、律师、律师事务所、法院、法官、立法机构、立法会议等相关主体和事件类信息资源的链接渠道。虽然部分资源通过超文本链接方式建立了信息单元间链接关系, 但该类链接属于单向链接, 无法实现反向跳转, 更无法实现间接链接关系的自动跳转。同时, 超链接自身无法承载语义内容, 通过超链接建立的关联无法表达数据间丰富多样的语义关联关系。

第三, 数据分散存储在不同数据源中, 使用不同的元数据规范描述, 以不同的格式呈现, 很难实现元数据映射和互操作, 数据整合和共享困难。例如, 中央政府网站发布的法律法规与最高人民法院网站发布的裁判文书分别存储在不同服务器中, 且这些数据源使用不同关系数据库模式和元数据方案, 很难实现这些数据的共享和关联。此外, 文件系统或关系型数据库数据存储方式, 是一种闭合的数据模型, 不利于元数据规范的扩展, 更不利于数据关联。

关联数据是万维网联盟(World Wide Web Consortium, W3C)推广的全新资源组织方式, 信息组织以数据为单位, 采用模式语言和本体语言形式化地描述元数据规范, 便于数据整合和机器读取。通过语义网技术体系构建的关联数据集, 能够提供全新数据利用方式,

可以实现图模式检索、社交网络分析和语义推理等。总之,语义网通过上述特征和优势,既提供了解决上述问题的新路径,也提供了数据利用和发掘的新机遇。

2 研究现状

针对如何使用语义网技术、依据关联数据原则组织和关联法律资源,并最终生成关联法律数据(linked legal data),国外已经展开了较多的研究和实践。

欧盟标准化委员会于2006年发布metalex本体,该本体通过形式化、语义化方式描述法律文献资源的元数据信息。阿姆斯特丹大学莱布尼茨法律中心的研究人员使用metalex本体,将荷兰政府开放数据门户“wetten.overheid.nl”中的法律法规转化为关联法律数据,并在Metalex Document Server网站发布^[1]。英国国家档案馆开发的网站,综合使用metalex、frbr和dct本体,以关联数据形式发布法律法规^[2];此外,该网站还将法律法规按照篇章结构进行细化切分,生成粒度精细的信息单元^[3]。芬兰阿尔托大学语义计算研究组的研究人员将Finlex Data Bank中的部分法律法规和裁判文书转换成关联数据格式,并在Semantic Finlex网站发布^[4]。意大利国家研究院法律信息理论与技术研究中心将法学领域的学术类文献资源,与法律法规、判例、出版机构相关联,形成法律数据云^[5]。布拉格查理大学的研究人员对metalex本体进行扩展,形成法律本体LEX后,将捷克的法律法规、裁判文书等发布为关联法律数据^[6]。康奈尔大学法学院法律信息研究所的研究人员正在开展关联法律数据项目(linked legal data project)^[7],该项目依据关联开放数据原则,使用信息抽取和自然语言处理技术,对《美国联邦法典》的文本进行处理,生成RDF元组数据库^[8]。

国内研究多数围绕法律本体展开。如贾君枝等研究如何建立法律本体与顶层本体的映射关系^[9];黄都培^[10]和钟振鸿^[11]等提出使用法律本体作为标引手段,实现对传统关键词检索功能的语义扩充;卢明纯^[12]和何庆^[13]等提出通过法律本体构建法律知识库,运用逻辑规则实现法律推理;张妮等设计了一个描述法律文本和裁判文书的法律本体,建立法律本体检索模型^[14]。虽然以上研究有助于实现法律信息语义检索和法律推理,但没有涉及运用关联数据原则组织和关联各类法律信息资源。张健文等研究了裁判文书元数据的抽取,并通过RDF表达,但没有涉及法律法规,更没有实现与其他类型法

律信息资源的关联^[15]。

综上所述,国外研究倾向于运用语义网技术、依据关联数据原则将已有法律资源转换成关联数据,并在Web上发布;国内研究倾向于构建法律本体以支持检索,缺乏运用关联数据关联已有法律资源的研究。因此,本文从法理学视角识别我国法学领域的各类法律资源,并依据语义网和关联数据原则设计法律资源的语义模型,以期对国内法律资源整合和关联的研究提供参考借鉴。

3 法律资源分析

界定法律资源范畴并依据一定标准对其分类,是法律资源组织和关联的前提。目前,从信息资源组织角度对我国法律资源的分类还没有公认的标准。在语义网体系内,资源可以代表任何事物;因此,法律资源也应该代表法律领域的任何事物。本文提出依据法理学视角来界定法律资源范畴,并设计其分类体系。

法理学以“一般法”为研究对象,涵盖法的整个领域,囊括全部法律事实,涉及法律运行的全过程。从该意义上讲,法理学的研究对象和范畴,可以在很大程度上代表信息组织意义上法律资源的分布范围。目前,我国法理学研究者主要从“法是什么”“法的结构”和“法的运行”等核心主题,来勾勒“法”的轮廓,分析其构成要素,描述和解释“法”这一复杂、多面的社会现象。“法是什么”,回答法的本质和特征等问题;“法的结构”,回答法的构成要素、法律渊源与效力等级、法律体系与法律部门、法律权利、法律义务和法律责任、法律关系、法律关系的主体和客体以及法律行为等问题;“法的运行”,回答立法、司法、执法、守法、法律监督,法律解释,法律推理和法律程序等问题。这些内容构成法的基本范畴。

针对上述法理学界定的范畴,还需进一步从信息组织角度,识别可以被描述和组织的资源类型。法律概念表现为自然语言的词汇和术语,这与信息科学领域的知识组织系统(Knowledge Organization System, KOS)中的“概念”是一致的,因此法律概念可以作为法律资源。在法律运行过程中,存在立法、司法、执法和守法等活动,这些活动可以视为法律事件,也可以作为法律资源来描述。法律规则和法律原则构成法律规范。法律规范通过语言文字表达,表现为规范性法律文件,而其载体是图情领域使用的法律文献概念,因此法律文献

可以作为法律资源。从法律关系角度看, 存在法律关系的主体和客体。法律主体通常为人或组织, 法律客体通常为物、智力成果、行为等, 这两类资源也可以作为独立的法律资源来描述。综合以上分析, 本文将法律资源进一步细分为法律概念、法律事件、法律文献、法律主体以及法律客体等资源。

4 法律资源语义模型

通过对各类法律资源的分析和描述, 可以形成各类法律资源的语义描述框架, 依据各类法律资源间的关联关系, 建立资源间链接, 进而形成各类法律资源的语义模型。在资源描述框架中, 本文综合使用frbr、metalex、dct等本体, 根据需求选择其中的类和属性词汇, 以综合利用不同本体在语义表达力上的优势, 既便于实现本体间的语义互操作, 也便于将生成的关联数据集与使用这些本体描述的其他关联数据集进行整合和关联。但上述本体无法表达我国法律制度的特殊情况, 因此本文在使用上述本体的基础上, 根据我国法律制度的特殊情况, 构建一个符合我国法律资源特色的本体, 将其命名为“linkedlaw”法律本体, 命名空间为“http://www.linkedlaw.org/”, 前缀为“ll:”。

4.1 法律概念资源

W3C将概念定义为思想的单元, 是人脑中的抽象实体, 独立于表达概念的具体词汇, 但还需要通过文字符号来表达, 即标签(label)^[16]。从信息组织角度而言, 法律概念可以来自法学领域专家构建的词汇集或分类词表, 其他类型法律资源的名称、别名或标题等, 以及对法律文献资源的文本进行分词形成的词组或短语。为区别法律概念与其他类型法律资源, 将法律概念资源放在命名空间“http://www.linkedlaw.org/concept/”中, 前缀为“llc:”。

对于法律概念资源, 本文没有设计新的模型来表达其内部关系, 而是完全使用简单知识组织系统(Simple Knowledge Organization System, SKOS)模型。W3C创建SKOS的最主要目的是支持叙词表、分类词表和主题词表等KOS的应用^[17]。而不同词表主要用于对各类信息资源进行标引, 形成各类信息资源的主题。本文是在KOS意义上建模, 并组织法律概念资源, 使用SKOS模型表达法律概念间的上位词、下位词、同义词

等语义关系, 利用法律概念资源组织和标引其他类型的法律资源。这样既能使法律概念成为整合和关联其他类型法律资源的重要节点和纽带, 又能使其作为其他类型资源集合中各实例的检索或浏览入口, 使用户能够以法律概念为中介对其他类型资源进行浏览或检索。

4.2 法律事件资源

法律运行包括立法事件、司法事件、执法事件和守法事件。立法事件是由立法主体制定、修改、废止法律的活动, 其结果包括创建新的法律和修改现有法律; 司法事件是人民法院的司法审判和检察机关的法律监督活动; 执法事件是执法主体实施和使用法律的活动, 多指行政执法; 守法事件是任何公民和组织遵守法律、享受权利、履行义务的活动。

针对法律事件资源, 综合使用sem^[18]、metalex和linkedlaw本体, 借助“sem:Event”和“metalex:Event”词汇从一般意义上描述事件类型。在linkedlaw中新增类词汇描述我国法律领域对法律事件的进一步细分。在“ll:法律事件”中设子类“ll:立法事件”“ll:司法事件”“ll:执法事件”和“ll:守法事件”, “ll:立法事件”下再设子类“ll:制定法律事件”“ll:修改法律事件”“ll:法律解释事件”和“ll:司法解释事件”。

4.3 法律文献资源

法律文献经常在法律信息科学这一新兴学科领域被提及和探讨, 目前还没有统一的定义^[19]。一般认为, 法律文献包括规范性法律文件、裁判文书和法学书刊。规范性法律文件主要是立法主体发布的各种具有法律效力、对社会活动具有规范作用的书面文件, 本文简称法律文件。从法律渊源角度来看, 法律文件可进一步分为宪法、法律、行政法规、地方性法规、自治条例、单行条例、国务院部门规章、立法解释、司法解释和实施细则等^[20]; 从法律体系角度来看, 法律文件可进一步分为宪法相关法、民商法、行政法、经济法、社会法、刑法、诉讼与非诉讼程序法等^[21]。此外, 法律文件在制定、修改和废止等活动中产生多个版本, 还可以分编、章、节、条、款、项、目, 形成树形结构体系。其中“条”按照整数递增的顺序贯穿文本始终^[22]。法律文件存在发布时间、生效时间、修订时间、废止时间、空间效力范围和语言文字等属性。在司法实践中, 人民法院越来越重视司法

案例的参考和指导作用。因此,裁判文书也是一类重要的法律文献资源。法学书刊主要包括围绕法学这一社会科学的研究和发展所产生的各种法学理论图书、期刊文章、会议论文和法学评论等。

针对法律文献资源,综合使用frbr、metalex和dct本体,以及本文构建的linkedlaw法律本体进行描述。以法律文件为例,使用“frbr:Work”描述法律体系中的某一部法律,如《著作权法》;使用“frbr:Expression”描述某一部法律的不同版本,如《著作权法》的1991年版、2001年版和2010年版;使用“frbr:Manifestation”描述某一具体法律版本以不同形式、载体和格式出版的具体实现形式;使用“frbr:realizationOf”描述“Expression”对“Work”的“实现”关系;使用“frbr:embodimentOf”描述“Manifestation”对“Expression”的“化身”关系。

参照国际图书馆协会联合会的定义并结合我国法律实践^[23],本文设定:“Work”仅代表“整体”,不划分“部分”,与其相关数据都是从整体描述其属性的元数据;“Expression”有“整体”和“部分”之分,将法律法规进一步划分为目录、编、章、节、条等子部结构,使每个子部均可以作为独立资源被访问获取,其中“条”是最小粒度资源。使用“metalex:fragmentOf”描述在法律规范的“Expression”层次上部分和整体的从属关系;使用“dct:title”“dct:identifier”描述标题、标识符等信息。本文在“linkedlaw”中增加新的类词汇。在“ll:法律文献”中设子类“ll:法律文件”“ll:裁判文书”和“ll:法学书刊”。在“ll:法律文件”下设子类“ll:宪法”“ll:法律”“ll:行政法规”等描述法律渊源,设子类“ll:宪法相关法”“ll:民商法”“ll:行政法”等描述法律部门。增加对象属性词汇描述文献资源间关系。用“ll:上位法”和“ll:下位法”描述法律间的效力层级关系,“ll:引用”和“ll:被引”描述法律文献间的引用关系,“ll:解释”和“ll:解释于”描述法律及其相应的法律解释、司法解释间的关系,“ll:细化”和“ll:细化于”描述实施细则对法律法规的细化关系。增加数值属性词汇描述文献的时间、空间、语言文字等属性。

在linkedlaw设计过程中,本文遵循设计规范,大部分对象属性都有其对应的逆属性,用“owl:inverseOf”描述(如“ll:引用 owl:inverseOf ll:被引.”),这样可实现两类法律资源间关系的互逆表达。

4.4 法律主体资源

围绕法律的运行,存在四类参与主体:立法主体、司法主体、执法主体和守法主体。立法主体负责制定和发布法律,包括国家权力机关和国家行政机关等;司法主体是依据法律拥有司法权的国家机关,包括人民法院和人民检察院,其中人民法院还可进一步按照级别细分;执法主体负责将法律运用到国家和社会的公共事务中,包括国家各级行政机关和法律授权的组织;我国法律规定,一切社会组织和个人都是守法主体,从这个意义上讲,立法主体、司法主体和执法主体都是守法主体的子类。需要强调的是,我国法理学中使用的主体概念,主要指作为组织的各类国家机关。本文从信息组织角度,认为各类主体还应包括各类机关中的公职人员,因此每类主体又进一步分为机关和人员。

针对法律主体资源,综合使用foaf和linkedlaw本体描述。使用“foaf:Organization”描述法人和其他组织,“foaf:Person”描述自然人,“foaf:name”“foaf:homepage”等描述名称、主页等相关信息。在linkedlaw中新增的类词汇,描述我国法学领域对主体类型资源的进一步细分。在“ll:法律主体”中设子类“ll:守法主体”,在“ll:守法主体”下设子类“ll:立法主体”“ll:司法主体”和“ll:执法主体”。

4.5 法律客体资源

在特定法律关系中,法律主体的权利和义务共同指向的对象是法律关系的客体,又称法律客体^[24]。我国法理学界对法律客体的分类存在歧义。结合众多学者思想,本文从信息组织角度,将法律客体分为有形物产、无形物产、行为和人身权利。其中,有形物产主要指法律客体意义上的“物”,包括各种自然有形物、人造有形物;无形物产主要指数字产品、智力成果,如软件、音像制品、著作权、商标权、专利权等;行为主要指法律主体具有法律意义的作为和不作为;人身权利主要包括人身自由、人格尊严、名誉权、肖像权等。

针对法律客体资源,综合使用schema和linkedlaw本体描述。使用“schema:name”描述客体名称,“schema:description”描述客体简介,“schema:image”描述客体图片。在linkedlaw中新增的类词汇,描述本文对我国法律客体设计的分类。在“ll:法律客体”中设子类“ll:有形物产”“ll:无形物产”“ll:行为”和“ll:人身权

利”,在“II:有形物产”中设子类“II:动产”和“II:不动产”。

4.6 链接各类法律资源

依据各类资源间的关联关系,在各类资源间建立链接,形成一张多层次、多粒度、多类型、全方位的关联法律数据网。为描述这种关联关系,在linkedlaw中增加多个对象属性词汇。

法律概念可以用来标引其他类型的法律资源。使用SKOS模型组织法律概念,发挥法律概念作为传统KOS的作用,使用这些法律概念标引其他类型的法律资源,由此形成法律概念与其他类型法律资源的标引或主题关系。在linkedlaw中加入一对互逆属性“II:标引”和“II:主题”,用“II:标引”描述法律概念对其他类型法律资源的标引关系,用“II:主题”描述其他类型法律资源以法律概念作为主题的关系。例如,《中国图书馆分类法(第五版)》中“D92”的概念“IIc:知识产权法”可以用来标引文献资源《著作权法》,对《商标法》第六十一条文本进行分词处理获得的“IIc:司法机关”可以用来标引各级人民法院实例资源。

法律文献可以作为法律事件的输入和输出。例如,将旧版本法律定义为法律修改活动的输入,由其产生的新版本法律则可定义为该法律修改活动的输出。使用互逆属性“II:使用”和“II:用于”描述法律文献作为法律事件的输入,使用互逆属性“II:产生”和“II:产生于”描述法律文献作为法律事件的输出。

法律文件规范法律主体和法律客体,使用互逆属性“II:规范”和“II:规范于”描述。裁判文书会提及相关的法律主体和法律客体,使用互逆属性“II:提及”和“II:提及于”描述。按照我国法理学界定义,法律客体是在特定法律关系中法律主体的权利和义务共同指向的对象。而法律关系中既存在权利主体,又存在义务主体。因此,在同一法律关系中一方法律主体对某一法律客体拥有权利的同时,同一法律关系中的另一方法律主体则对该法律客体承担义务。使用互逆属性“II:权利”和“II:权利人”描述拥有权利的法律主体和法律客体间的法律关系,使用互逆属性“II:义务”和“II:义务人”描述承担义务的法律主体和法律客体间的法律关系。

通过上述操作,在各类法律资源间建立链接关系,形成法律资源的语义模型(见图1),据此可以构建整合和关联各类法律资源的关联法律数据网。

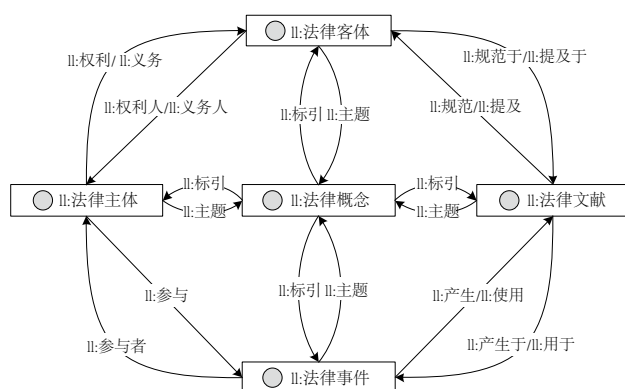


图1 法律资源语义模型

5 关联法律数据构建实例

本文通过原始法律文本数据,对提出的法律资源语义模型构建关联法律数据过程进行实验。本文在中国裁判文书网获取一份裁判文书,该裁判文书是苹果公司与北京磨铁数盟信息技术有限公司对《明朝那些事儿》的信息网络传播权的侵权纠纷二审结果。通过分析原始文本可知:北京市高级人民法院作为司法主体,参与了一项司法事件,可按以下描述。

```
<foaf:Organization rdf:about="http://www.linkedlaw.org/北京市高级人民法院">
  <rdf:type rdf:resource="http://www.linkedlaw.org/高级人民法院"/>
  <foaf:homepage rdf:resource="http://bjgy.chinacourt.org/">
  <rdfs:label>北京市高级人民法院</rdfs:label>
  <II:参与 rdf:resource="http://www.linkedlaw.org/苹果公司与北京磨铁数盟信息技术有限公司著作权权属侵权纠纷二审"/>
</foaf:Organization>
```

该二审审判事件,有多个参与者,产生裁判文书,可按以下描述。

```
<sem:Event rdf:about="http://www.linkedlaw.org/苹果公司与北京磨铁数盟信息技术有限公司著作权权属侵权纠纷二审">
  <rdf:type rdf:resource="http://www.linkedlaw.org/司法事件"/>
  <II:参与者 rdf:resource="http://www.linkedlaw.org/苹果公司"/>
  <II:参与者 rdf:resource="http://www.linkedlaw.org/磨铁公司"/>
  <II:参与者 rdf:resource="http://www.linkedlaw.org/北京市高级人民法院"/>
```

<ll:产生 rdf:resource="http://www.linkedlaw.org/苹果公司与北京磨铁数盟信息技术有限公司著作权权属侵权纠纷二审民事判决书/2014/zh"/>

<rdfs:label>苹果公司与北京磨铁数盟信息技术有限公司著作权权属、侵权纠纷二审</rdfs:label>
</sem:Event>

该裁判文书提及法律客体、司法主体，引用法律条款，可按以下描述。

<frbr:Expression rdf:about="http://www.linkedlaw.org/苹果公司与北京磨铁数盟信息技术有限公司著作权权属侵权纠纷二审民事判决书/2014/zh"/>

<rdf:type rdf:resource="http://www.linkedlaw.org/裁判文书"/>

<rdfs:label>苹果公司与北京磨铁数盟信息技术有限公司著作权权属侵权纠纷二审民事判决书（2014）</rdfs:label>

<ll:提及 rdf:resource="http://www.linkedlaw.org/明朝那些事儿的信息网络传播权"/>

<ll:提及 rdf:resource="http://www.linkedlaw.org/北京市第二中级人民法院"/>

<ll:引用 rdf:resource="http://www.linkedlaw.org/著作权法/2010/zh/10"/>
</frbr:Expression>

通过分析著作权法中的文本可知，1990年作为立法主体的“全国人民代表大会”在立法事件“第七届全国人民代表大会常务委员会第十五次会议”中首次制定《著作权法》，按照本文设计的法律文献资源描述框架，该次立法事件既产生了一个“frbr:Work”层次的《著作权法》，同时又产生了“frbr:Expression”层次的《著作权法（1991年版）》，该立法事件可以进一步认定为“制定法律事件”，可按以下描述。

<sem:Event rdf:about="http://www.linkedlaw.org/第七届全国人民代表大会常务委员会第十五次会议">

<rdf:type rdf:resource="http://www.linkedlaw.org/立法事件"/>

<rdf:type rdf:resource="http://www.linkedlaw.org/制定法律事件"/>

<ll:参与者 rdf:resource="http://www.linkedlaw.org/全国人民代表大会常务委员会"/>

<ll:产生 rdf:resource="http://www.linkedlaw.org/著作权法"/>

<ll:产生 rdf:resource="http://www.linkedlaw.org/著作权法/1991/zh"/>

<rdfs:label>第七届全国人民代表大会常务委员会第十五次会议</rdfs:label>

<sem:hasTime rdf:resource="http://www.linkedlaw.org/date/1990-09-07"/>
</sem:Event>

同时，《著作权法》的版本信息可按以下描述。

<frbr:Work rdf:about="http://www.linkedlaw.org/著作权法">
<rdf:type rdf:resource="http://www.linkedlaw.org/法律"/>

<rdf:type rdf:resource="http://www.linkedlaw.org/民法商法"/>

<rdfs:label>中华人民共和国著作权法</rdfs:label>

<frbr:realization rdf:resource="http://www.linkedlaw.org/著作权法/1991/zh"/>

<frbr:realization rdf:resource="http://www.linkedlaw.org/著作权法/2001/zh"/>

<frbr:realization rdf:resource="http://www.linkedlaw.org/著作权法/2010/zh"/>

</frbr:Work>

<frbr:Expression rdf:about="http://www.linkedlaw.org/著作权法/2010/zh">

<rdf:type rdf:resource="http://www.linkedlaw.org/法律文件"/>

<rdf:type rdf:resource="http://www.linkedlaw.org/法律"/>

<rdf:type rdf:resource="http://www.linkedlaw.org/民法商法"/>

<rdfs:label>中华人民共和国著作权法（2010）</rdfs:label>

<frbr:realizationOf rdf:resource="http://www.linkedlaw.org/著作权法"/>
</frbr:Expression>

《著作权法》（2010年版）第十条，规范作为法律客体类型的信息网络传播权和作为法律主体类型的著作权人，对其文本进行分词操作，还可以获取法律概念，可按以下描述。

<frbr:Expression rdf:about="http://www.linkedlaw.org/著作权法/2010/zh/10">

<rdfs:label>中华人民共和国著作权法（2010）第十条</rdfs:label>

<metalex:fragmentOf rdf:resource="http://www.linkedlaw.org/著作权法/2010/zh"/>

<ll:主题rdf:resource="http://www.linkedlaw.org/concept/著作权"/>

<ll:主题rdf:resource="http://www.linkedlaw.org/concept/信息网络传播权"/>

</frbr:Expression>

<ll:客体 rdf:about="http://www.linkedlaw.org/明朝那些事儿的信息网络传播权">

<rdf:type rdf:resource="http://www.linkedlaw.org/无形物产"/>

<rdfs:label>《明朝那些事儿》的信息网络传播权</rdfs:label>

<ll:规范于 rdf:resource="http://www.linkedlaw.org/著作权法/2010/zh/10"/>

</ll:客体>

<foaf:Organization rdf:about="http://www.linkedlaw.org/磨铁公司">

```

<rdf:type rdf:resource="http://www.linkedlaw.org/守
法主体"/>
<rdfs:label>北京磨铁数盟信息技术有限公司</rdfs:label>
<ll:规范于 rdf:resource="http://www.linkedlaw.org/
著作权法/2010/zh/10"/>
</foaf:Organization>

```

法律概念可以再作为其他法律文件的主题,可按以下描述。

```

<skos:Concept rdf:about="http://www.linkedlaw.org/
concept/信息网络传播权">
  <skos:prefLabel xml:lang="zh">信息网络传播权</
skos:prefLabel>
  <skos:broader rdf:resource="http://www.linkedlaw.
org/concept/著作权">
</skos:Concept>
<frbr:Expression rdf:about="http://www.linkedlaw.org/信
息网络传播权保护条例/2013/zh">
  <rdfs:label>信息网络传播权保护条例 (2013) </rdfs:label>
  <ll:主题 rdf:resource="http://www.linkedlaw.org/
concept/信息网络传播权"/>
</frbr:Expression>

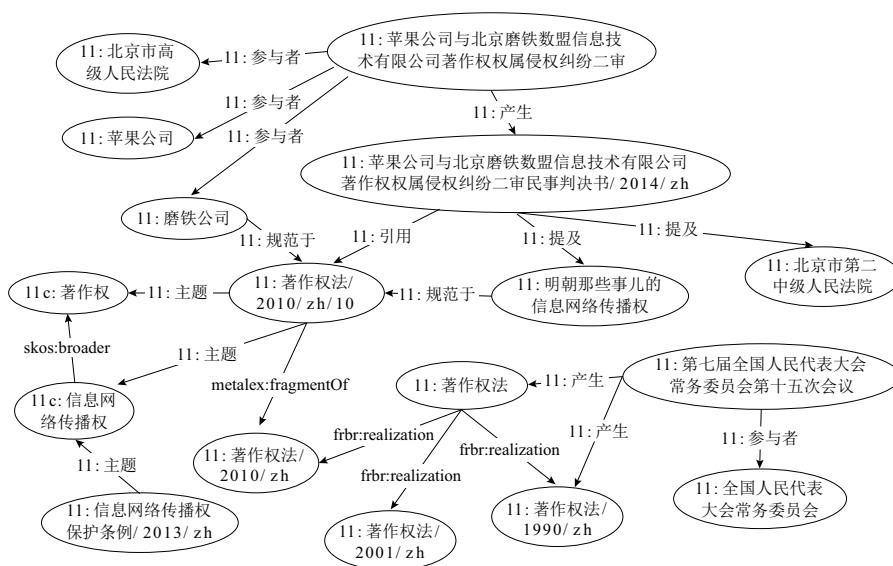
```

通过上述资源描述操作,可以实现我国法学领域内各类型法律资源关联,形成关联法律数据网(见图2)。通过该关联法律数据网,用户可以在各类法律资源间进行导航和查询,获取相关信息,更好地满足其法律信息需求。

6 结语

本文首先从法理学视角识别我国法律资源类型,从信息组织角度识别实体属性和关系;其次,依据语义网技术体系标准,分别对各类信息资源分析和描述;再次,依据各类信息资源间关联关系,建立信息资源链接;最后,构建法律资源语义模型,并通过原始文本数据对关联法律数据的构建过程进行实验。

本文还存在一定不足,如没有实现从法律文本中自动识别实体、提取特征并建立链接,没有更深入地研究大规模关联数据集的挖掘和可视化等,后续将对此进一步展开研究。



参考文献

[1] HOEKSTRA R. The MetaLex document server: legal documents as versioned linked data[J/OL]. The Semantic Web-ISWC 2011, 2011, 7032: 128-143 [2017-01-10]. http://link.springer.com/chapter/10.1007/978-3-642-25093-4_9. DOI: 10.1007/978-3-642-25093-4_9.

[2] SHERIDAN J. Legislation.gov.uk[EB/OL]. [2017-01-22]. <http://blog.law.cornell.edu/voxpath/tag/legal-linked-data/>.

[3] CARRASCO M H, MARTÍNEZ-GONZÁLEZ M M, DE LA FUENTE REDONDO P. Data models for version management of legislative documents[J/OL]. Journal of Information Science, 2013, 39(4): 557-572 (2013-02-13) [2017-02-10]. <http://journals.sagepub.com/doi/10.1177/016555>

1512473723. DOI: 10.1177/0165551512473723.
- [4] FROSTERUS M, TUOMINEN J, HYVÖNEN E. Facilitating re-use of legal data in Applications-Finnish law as a linked open data service[J]. Legal Knowledge and Information Systems, 2014: 115-124.
- [5] AGNOLONI T, FRANCESCONI E, SAGRI M T, et al. Linked data in the legal domain[J/OL]. [2017-03-10]. <http://www.cersi.it/itais2011/pdf/68.pdf>.
- [6] NEČASKÝ M, KNAP T, KLÍMEK J, et al. Linked open data for legislative domain-ontology and experimental data[J/OL]. Business Information Systems Workshops, 2013, 160: 172-183 [2017-02-10]. http://link.springer.com/chapter/10.1007/978-3-642-41687-3_17. DOI: 10.1007/978-3-642-41687-3_17.
- [7] Legal Information Institute. LII Announce[EB/OL]. (2012-04-06) [2017-02-20]. <https://blog.law.cornell.edu/blog/tag/linked-legal-data/>.
- [8] CASELLAS N, BRUCE T R, FRUG S S, et al. Linked legal data: improving access to regulations[C]//Proceedings of the 13th Annual International Conference on Digital Government Research. ACM, 2012: 280-281.
- [9] 贾君枝, 郭丹丹. 法律框架本体与顶层本体SUMO的映射研究[J]. 图书情报工作, 2008(4): 74-77, 103.
- [10] 黄都培. 法律信息语义检索方法研究[J]. 法律文献信息与研究, 2009(4): 1-10.
- [11] 钟振鸿, 印润远, 于庆梅. 基于本体驱动的法律信息检索模型[J]. 微计算机信息, 2007(30): 178-180.
- [12] 卢明纯. 基于OWL本体的法律知识库原型系统的设计和实现[J]. 现代情报, 2010(7): 34-38.
- [13] 何庆, 汤庸, 黄永钊. 基于本体的法律知识库的研究与实现[J]. 计算机科学, 2007(2): 175-177.
- [14] 张妮, 杨遂全, 蒲亦非. 我国法律本体检索模型的研究[J]. 法律方法, 2015(2): 104-117.
- [15] 张健文, 朱建林, 黄都培, 等. 一种基于RDF的法律判决书语义标注方法[C]//中国计算机用户协会信息系统分会2012年第22届信息交流大会论文集, 2012: 61-65.
- [16] W3C. SKOS simple knowledge organization system primer[EB/OL]. [2017-03-03]. <https://www.w3.org/TR/2009/NOTE-skos-primer-20090818/>.
- [17] W3C. SKOS simple knowledge organization system reference[EB/OL]. [2017-03-03]. <https://www.w3.org/TR/2009/REC-skos-reference-20090818/>.
- [18] HAGE W R V, MALAISÉ V, SEGERS R, et al. Design and use of the Simple Event Model(SEM)[J]. Web Semantics Science Services and Agents on the World Wide Web, 2011, 9(2): 128-136.
- [19] 田建设. 我国法律文献检索教材之检讨[J]. 法律文献信息与研究, 2006(3): 19-33.
- [20] 中华人民共和国立法法[EB/OL]. (2015-03-18) [2017-02-24]. http://www.npc.gov.cn/npc/dbdhy/12_3/2015-03/18/content_1930713.htm.
- [21] 国务院新闻办公室. 《中国特色社会主义法律体系》白皮书发布(全文)[EB/OL]. (2011-10-27) [2017-02-28]. http://www.gov.cn/jrz-zg/2011-10/27/content_1979498.htm.
- [22] 立法技术规范(试行)(一)及其说明[EB/OL]. (2011-04-01) [2017-02-24]. <http://www.ccpc.cq.cn/Home/Index/more/id/63962.html>.
- [23] IFLA Study Group on the functional requirements for bibliographic records. Functional requirements for bibliographic records final report [R/OL]. [2017-03-27]. https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf.
- [24] 严存生. 法理学[M]. 北京: 中国政法大学出版社, 2009: 197.

作者简介

邢启迪, 男, 1982年生, 硕士研究生, 研究方向: 语义网、信息和知识组织, E-mail: xingqidi@mail.bnu.edu.cn。
耿骞, 男, 1965年生, 博士, 教授, 研究方向: 信息检索、网络信息管理、管理信息系统, E-mail: gengqian@bnu.edu.cn。
赵盼云, 女, 1992年生, 硕士研究生, 研究方向: 政府信息管理、关联数据, E-mail: flyzpy@qq.com。
景然, 女, 1992年生, 硕士研究生, 研究方向: 政府信息管理、数据挖掘、信息检索, E-mail: jingran@mail.bnu.edu.cn。

Research on Building Semantic Model for Legal Resources

XING QiDi, GENG Qian, ZHAO PanYun, JING Ran
(School of Government, Beijing Normal University, Beijing 100875, China)

Abstract: First, from the jurisprudence perspective, this paper identifies the types of resources in China legal domain. Second, from the information organizing perspective, the paper analyzes each type of legal resources, identifies properties and relationships. Third, the paper reuses frbr, metalex and dct ontologies and constructs a new legal ontology to design description framework for all types of legal resources. Finally, according to the semantic relationships among all types of legal resources, the paper establishes linkages among these resources and builds the legal resource semantic model. The demo shows that the semantic model can be used to generate a web of linked legal data which spans multiple levels, multiple granularities, multiple types of legal resources. The linked legal data can meet legal information needs better.

Keywords: Legal Information Organization; Linked Legal Data; Legal Ontology

(收稿日期: 2017-04-27)