

# 法律人工智能新范式：封闭与开放的二元兼容

余成峰\*

**摘要** 符号主义与联结主义两种经典人工智能范式，以及由此形成的基于规则和基于数据的法律推理模型，对应于规范封闭与认知开放这一古老的法理学命题。符号主义强调规范的封闭，联结主义强调认知的开放，两条路径代表了两种迥异的技术范式，在其背后潜藏着深刻的哲学思想差异，并导向两种完全不同的法律发展模式。法律人工智能的封闭与开放，因此涵盖了一系列重要的二元论法理学命题。可以围绕这组二元概念，在形式理性法和实质理性法的思想传统中，借助系统论法学的概念工具，重构人工智能法学的核心理论命题。大数据智能技术的单向驱动，将在法律的社会、时间和事实面向引发深刻的法律变异。这要求从法律的功能、领域、过程和结构层面，重新定位法律人工智能的三种演化类型。应当从规范封闭与认知开放的二元兼容视角，发展更为复杂的法律理论，以此设计法律人工智能的发展路线图，引导数字时代智能法律的良性发展。

**关键词** 法律人工智能 规范封闭 认知开放 个性化法律 反思法

## 引言

20世纪以来，人工智能的发展进程呈现出符号主义(symbolicism)与联结主义(connectionism)两条典型路径。相应地，在人工智能的法律应用情景中，即本文所称的“法律人工智能”领域，形成了基于规则的推理模型(rule-based reasoning)和基于数据的推理模型(data-based

\* 北京航空航天大学人文与社会科学高等研究院副教授。本文系国家社科基金重点项目“系统论视野下的数字法治基本问题研究”(项目编号:22AZD149)的阶段性研究成果。

reasoning)两种范式。<sup>〔1〕</sup>符号主义强调规则上的封闭性,联结主义强调数据上的开放性。这两条不同的法律人工智能范式,恰恰对应于法律规范的封闭性与开放性这一二元论法理学命题。<sup>〔2〕</sup>

观察我国目前关于人工智能的法学研究,可发现各持一端或不加区分这两种技术路径及范式而讨论法律人工智能问题的倾向。<sup>〔3〕</sup>然而,我们需要认识到:这两条路径代表了两种迥异的人工智能技术范式,不同的技术路径背后潜藏着深刻的哲学思想差异,并影响到法律人工智能的不同法律模式及发展方向,包括对法律人工智能的性质和功能应如何理解、对人工智能相关法律的制定、解释、适用和实施活动应如何展开等重大问题。

质言之,法律的人工智能化,有可能将法律引向两条完全不同的道路,将法律固有的封闭性和开放性二元张力推向极致。符号主义基于规则的封闭运作,强调形式主义的演绎推理;联结主义基于数据的开放认知,强调实质主义的统计推理。法律人工智能应当优先适用哪一条路径而发展,这在根本上需要面对“法律是什么”以及如何在法理学上协调规范封闭与认知开放的二元张力问题。这又牵涉到对法律的功能特征、推理模式和运作逻辑的重新理解,并由此更好地设计法律人工智能发展路线图的问题。

在数字化智能化技术驱动下,法律的封闭性与开放性这组二元张力是否会向各自的极限化方向发展,或是出现一边倒的失衡?两种技术范式将会齐头并进,取长补短,还是持续摇摆、相互冲突?这是人工智能法学研究亟需思考和应对的核心命题。这要求我们从法律的功能、领域、过程和结构等层面,重新思考法律人工智能的三种演化类型,安顿智能法律的规范封闭与认知开放这组二元张力,从法律人工智能自我反思宪治体系的建构视角,探寻人工智能法律的发展方向。

## 一、法律人工智能的两种范式及其思想基础

### (一)符号主义:基于规则的法律推理模型

#### 1. 基本思想

在符号主义视角下,智能是通过“物理符号系统”进行信息处理的模式。法律智能的搭建

---

〔1〕 有关符号主义与联结主义的经典文献,see Margaret Boden, *Mind As Machine: A History of Cognitive Science*, Oxford: Oxford University Press, 2008, pp. 701-1001. 在法律人工智能领域,对应符号主义与联结主义,分别形成基于规则的推理模型(基于案例的推理模型同属符号主义的分支)和基于数据的推理模型两种范式。晚近的机器学习,包括法律大语言模型,都从属联结主义的谱系。有关法律人工智能推理模型的经典研究,可参见(美)凯文·阿什利:《人工智能与法律解析》,邱昭继译,商务印书馆2020年版。

〔2〕 法律规范的封闭性和开放性,涵盖了一系列重要的二元论法理学命题。例如法律的统一性/灵活性、安定性/可变性、唯理主义/经验主义、规范性期望/认知性期望、演绎推理/归纳推理、义务论/后果主义、不学习/学习、教义法学/社科法学、自治型法/回应型法、形式法治/实质法治等等。

〔3〕 刘艳红同样批评人工智能法学研究存在明显的概念附会现象,参见刘艳红:“人工智能法学研究的反智能化批判”,《东方法学》2019年第5期,第119—126页。

首先需要打造一套“符号系统”,以实现法律图灵机的自动计算。将复杂的法律认知过程还原为简单的机械化操作,按照预装知识库的规则对输入的法律事实进行计算化处理。其基本预设是,形式化的法律符号系统可以有效标示法律世界的真实对象。智能机器通过三段论的逻辑推理,可以实现符号的封闭化操作,实现自动化的形式演绎。此即司马贺和纽艾尔的“通用问题求解器”(general problem solver)思想。<sup>〔4〕</sup>这一思想体系有以下几方面的核心特征。

第一,符号主义的哲学基础是一种功能主义(functionalism)立场。按照万能图灵机的理念,作为一种可以运算任何合法运算式的抽象计算模型,法律智能可以被有效从人类大脑迁移到机器演算,无非是将符号化的封闭操作过程实现于不同的物理基础之上。功能具有多重可实现性(multiple realizability),在功能之间具有等价性(functional equivalent)。

第二,符号主义预设了一种外在语义论思想。外在主义的语义学理论(externalist semantic theory)认为:任何知识表征的语义内容,都指涉表征符号与其表征的外部对象之间的关系。在这一视角下,所有法律命题都可以被有效映射为统一的形式符号,这些形式符号一一对应于外在的法律世界,因此是客观法律世界的语言图像。对法律的形式符号进行封闭操作,即可等价于对开放法律世界的把握。

第三,符号主义内含了莱布尼茨式的“理想语言”(characteristica universalis)思想:认为法律语言是由符号以及联系符号的逻辑构成,诸符号的所指清晰明确,而相关的逻辑可以通过计算化的方式演绎。<sup>〔5〕</sup>一切可能的自然法律语言,都存在着某些始终有效的逻辑法则,存在着数量有限的形式参数,法律的“普遍语法”具有抽象的形式性和完美性,由它们来决定自然法律语言的表层句法形式。法律的人工智能化,其关键就是找到可以贯穿一切法律语言的先天原则和参数,即将自然法律语言化约种种经验之后的形式“剩余物”。在自然法律语言和人工法律语言之间,可以通过这一理想法律语言进行桥接,将其作为转化的“中间语”(intermediate language),这一“中间语”具有容纳任何自然法律语言的潜在语义结构。这种“理想语言”也类似乔姆斯基的“深层语法”思想,<sup>〔6〕</sup>均是将其作为自然法律语言的深层抽象结构,以作为法律机器进行形式符号操作的基础。

第四,符号主义尝试建立关于外在法律世界的理想语义学模型,通过对该模型中法律原子命题的逻辑演绎,通过可计算的方式,对任何一个法律命题的含义进行自动解析。基于规则的推理建模的首要工作是法律知识表示,即将法条和案例知识表示为可计算的法律知识库,为自动法律推理提供基础。<sup>〔7〕</sup>在此,符号主义假定了外部的法律世界具有稳定的形而上学结构,

〔4〕 See Allen Newell and Herbert Simon, “Computer Science as Empirical Inquiry: Symbols and Search,” *Communications of the ACM*, Vol. 19, No. 3, 1976, pp. 113-126.

〔5〕 参见(德)莱布尼茨:《莱布尼茨自然哲学著作选》,祖庆年译,中国社会科学出版社1985年版,第8页。

〔6〕 See Noam Chomsky, “Deep Structure, Surface Structure, and Semantic Interpretation,” in Danny D. Steinberg and Leon A. Jakobovits (eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*, London: Cambridge University Press, 1971, pp. 183-216.

〔7〕 参见熊明辉:“法律人工智能的推理建模路径”,《求是学刊》2020年第6期,第92页。

符号之所以可以表征实在,是因为符号和实在共享相同的逻辑形式(例如法律的道义逻辑)。法律人工智能的核心任务,即在于通过符号表征对这一逻辑形式作出准确的刻画。

第五,符号主义的运作封闭理念,源自其“整体论”和“内在论”的思想。<sup>〔8〕</sup>根据整体论,法律符号的含义可以通过它在符号系统中的推理关系得以界定;根据内在论,法律符号的含义可以借由符号系统的语义网本身(无需指涉符号系统之外的事物)得以明确。因此,在法律人工智能的程序设计中,对于一个算子(operator)的界定,可以基于相关规则向量空间的映射实现,对于规则的描述,就会自动涉及该算子在法律语义网中的推理地位。基于规则的推理模型,是将法条表达为机器可理解的知识库,以供自动化的推理引擎随时调用。既然规则可以在符号层面表达,而计算机又可以自动执行演绎推理,那么,从计算层面建模法律推理过程就不存在根本障碍。<sup>〔9〕</sup>由此,就可以保证法律符号体系的无欠缺性,实现系统运作的封闭性。

第六,符号主义持有逻辑原子主义(logical atomism)思想,即认为人工智能法律所要表征的对象,是确定的、明晰的、离散的事物。真实的法律世界可以被界分为一系列的原子法律事实,这些法律事实可以通过在道义逻辑层面彼此独立的符号命题得到表征。<sup>〔10〕</sup>原子命题可以表征法律世界的原子事实,复合命题可以表征法律世界的复合事实,由此推演,封闭的法律符号系统就可以完整地表征整个法律世界,而无需任何认知上的开放。对于法律世界的描绘,可以分解为自身无法再分解的原子命题(例如霍菲尔德式命题)。只要法律命题的结构是按照纯粹形式逻辑组建的,那么任何外在的经验事实都与命题的规范结构无关。符号主义因此希望在纯粹的形式刻画中,排除掉所有经验杂质的干扰,以实现自动化的封闭运作。

第七,在符号主义视野下,智能的核心是对语义关系的把握能力,而这种语义能力指向主体的“意向性”(intentionality)。换言之,在语义符号的背后,预设了某种道德主体的存在,只有通过道德主体的能动性和意向性,才能将语义符号和符号表征的对象真正联系起来。因此,符号主义内含康德式道义论的伦理学传统,即将理性主体的道德能动性,作为道义逻辑演绎和法律责任归因的基础。在遵守义务和后果权衡之间,符号主义会果断排除认知层面的计算和实用权衡,以“绝对命令”为依据,实现形式化的演绎推理。

立足以上七个方面的思想理念,符号主义尝试将法律概念和法律命题全面符号化、形式化、数学化,超越自然法律语言,用一劳永逸的方式完成对自然法律世界的全面表征任务。通过系统公理的抽象设定,将法律推理全面计算化,将法律运作完全自动化,由此形成一个运作封闭的无逻辑矛盾的形式推理体系,使得任何一个法律命题都可以成为机器计算的演绎对象。由此,就可以将法律系统的封闭运作逻辑发挥到极致,摆脱环境开放认知的负担,超越人类法

〔8〕 这一思想的主要代表是费尔迪南·德·索绪尔(Ferdinand de Saussure),这种结构主义的符号观认为,意义是符号系统内部意指过程的结果,符号是自给自足的先验存在,所指也是先验存在的观念世界,参见(瑞士)费尔迪南·德·索绪尔:《普通语言学教程》,高名凯译,商务印书馆1980年版。

〔9〕 参见熊明辉,见前注〔7〕,第91页。

〔10〕 “我称自己的学说为逻辑原子主义的理由,是因为我想在分析中取得的作为分析中的最终剩余物的原子并非物质原子而是逻辑原子。”Bertrand Russell, *Logic and Knowledge*, Nottingham: Spokesman Books, 2007, p. 214.

律决策的偏见与错误,实现从“判决”到“算决”的跃迁。<sup>〔11〕</sup>

## 2. 符号主义的主要问题

首先,符号主义认为,法律人工智能的核心是“法律推理的逻辑表示问题”。<sup>〔12〕</sup> 这里其实隐含了一种全能理性的“上帝视角”:程序员能够获得关于法律世界的充分知识,这些专家知识可以作为系统的先天规则进行预存。而根据晚近的建构主义认识论,语义不是外在的,结构也不是先天与固定的,特定系统的运作过程在实际塑造系统的结构,不同的系统运作历史决定性地影响着符号结构的形态。<sup>〔13〕</sup> 换言之,是运作产生了结构,而不是结构决定着运作,不存在由先天符号结构单向决定法律系统运作的情况。法律系统的规范封闭性,并不由静态的符号规则保证,而是在动态的运作递归网络中产生,系统在复杂环境中的自主运作,逐渐“涌现”出相应的规则体系。<sup>〔14〕</sup> 正因如此,法律系统的规范封闭是以系统的认知开放作为前提,而不存在作出事先一劳永逸的符号刻画的可能,无论这种规则刻画在形式逻辑上多么完美,都需要法律系统在复杂环境中通过持续的运作进行修正和检验。

其次,目前在“基于规则的推理模型”中运用的算法工具,均是弗雷格式的一阶谓词逻辑的变种。<sup>〔15〕</sup> 这种基于真值函数的形式逻辑演绎,不可避免导致相关法律命题的语义学与语用学背景被全面忽视。而20世纪后半叶的语用学转向,已然打破了“形上世界”的假象,在语用学视角下,现代逻辑看似完美的语义表征,实际是对日常语言现象的严重误解。法律命题不具有凌驾所有语境的先天真值性,而只能是相关语境下的恰当用法。正如哈特所说,法律规则乃是一个具有确定的核心、模糊的边缘和开放的结构,众多法条都包含不确定的术语和概念,法律语义具有含混性与模糊性。<sup>〔16〕</sup> 质言之,法律概念只是法律沟通的媒介,其意义只能在法律的语用沟通过程中生成,并在持续的法律沟通中不断发生语义的演变。因此,不可能对所有关键的法律范畴和命题进行完美的语义表征,也无法假装系统已经拥有关于法律世界的完美知识,已然掌握法律命题的所有逻辑关系。法律的语用沟通过程具有面向未来的高度开放性和不确定性。

其三,法律系统的语义表征既需要稳定性和确定性,以实现规范封闭;又需要避免静态性和僵化性,以实现认知开放。如果符号无法灵活地表征现实,就难以“与时俱进”,灵活应对复

〔11〕 参见王涌:“道义逻辑、人工智能与法律——霍菲尔德法律关系形式理论的应用”,《经贸法律评论》2020年第2期,第80页。

〔12〕 参见熊明辉,见前注〔7〕,第90页。

〔13〕 有关建构主义认识论,可参见刘彬彬:《卢曼的社会系统论及其建构主义哲学基础》,华中科技大学2021年博士学位论文,第63页。

〔14〕 规则与标准的区分,以及任意规定、不确定法律概念、空白条款、一般性条款,都可作为在法律系统运作中实现“涌现”(emergence)的规则实例,see Louis Kaplow, “Rules versus Standards: An Economic Analysis,” *Duke Law Journal*, Vol. 42, No. 3, 1992, pp. 557-629.

〔15〕 戈特洛布·弗雷格(Friedrich Ludwig Gottlob Frege)从数学中的函数结构出发,构造了形式语言,使得逻辑推理表示为机械计算成为可能。参见(德)G. 弗雷格:《算术基础》,王路译,商务印书馆1998年版,第105—123页。

〔16〕 参见(英)哈特:《法律的概念》,张文显等译,中国大百科全书出版社1996年版,第123—127页。

杂环境；如果认知过于开放，法律系统则可能被外部世界的复杂性同化，从而丧失系统运作的自主性。而在当前技术条件下，智能法律的符号系统只能通过程序员预装，系统自身缺乏自主更新的能力，因而只是自动化而不是自主化的操作，充其量只是对人类法律表征符号的机械化执行。“迄今为止所有公理化的形式推理系统，都不允许后天的经验变化对先天输入系统的公理集的有效性产生任何影响。”〔17〕与此同时，由于法律事务的复杂性与开放性，法律知识库的不断扩展，又必然会带来“基本谓词膨胀”的问题，这会导致智能系统的运作变得愈加迟钝，系统对原有法律知识库进行修正的成本也会不断攀升。此外，符号主义也难以避免法律人工智能的“框架问题”：程序员必须从外部以“框架公理”的形式，人工输入“何规则与何规则不相关”的信息，否则系统就没有能力自主判断“何规则与何规则不相关”，而与此相关的框架公理不断叠床架屋，则会让智能系统陷入难堪重负的困境。〔18〕

法律“专家系统”尝试将法律专业知识用形式语言进行整编，然后利用演绎推理规则自动解析法律结论。但是，过于简化、缺乏弹性的形式符号无法应对不确定的法律世界。这种经人类“填鸭式教学”（手工输入知识）而成为“专家”的“智能系统”，由于其自身不具备自主学习能力而始终面临知识的瓶颈。〔19〕外部世界具有开放性和不确定性，如果法律系统没有能力在新案件与旧案件之间建立恰当的类比，它就会因为受限于既有的知识库而无法应对新情况。进言之，在特定的法律输入集与输出集之间建立涵摄关系的过程，需要构建庞大的法律公理集，这必然会带来巨大的编程问题，而如何从公理集中迅速调取合适的子集作为法律推理的起点，又会面临棘手的工程学难题。因此，智能系统若要成为真正的自主系统，其先天知识和后天知识的界限就需要是相对模糊的，因为系统不可能事先获知所有可能的法律事件类型。

## （二）联结主义：基于大数据的法律推理模型

### 1. 基本思想

针对符号主义的弊端，联结主义认为实现智能不止于符号主义一种路径，符号推理只是算法化的策略之一。法律算法的构建可以脱离抽象的公理演绎，变成一种分布式、局域化的计算过程。放弃形式化的规则表征，缩小系统知识库的规模，通过大量投喂法律数据进行训练，直接在海量数据中寻找法律的相关性模式。其基本工作原理，就是通过大量的人工标注数据训练，经由一个内部参数可以调整的大型人工神经网络系统的操作，形成特定法律输入与输出之间的稳定映射关系。法律大数据的深度学习，可以将具有特定特征的法律输入高效识别为法律语义网中的一些既定标签。由此，不需要自上而下地向系统输入公理化的形式符号规则，只需要自下而上地通过样本训练和数据挖掘，实现相关案件的自动搜索、识别和匹配，从而形成法律预测和案件处理的能力。

尤其是在法律事实和证据的认定上，大数据推理相比符号推理，可以直面经验世界的开放

〔17〕 徐英瑾：《哲学证成与机器编程》，复旦大学出版社2023年版，第96页。

〔18〕 参见夏永红、李建会：“人工智能的框架问题及其解决策略”，《自然辩证法研究》2018年第5期，第3—8页。

〔19〕 参见刘东亮：“新一代法律智能系统的逻辑推理和论证说理”，《中国法学》2022年第3期，第155页。

数据,且支持认知上的自我修正。<sup>[20]</sup>理想的状态是,一方面从系统外部输入大量开放数据,另一方面将处理结果映射为法律语义内容,交付给更高的系统符号层处理。通过运用多种统计学手段(深度学习、贝叶斯网络、遗传算法、模糊逻辑等),进行法律大数据的模式识别、案件匹配和概率预测,进行证据认定、法律检索、类案推送、量刑辅助,从而实现从“法律审判”到“法律预测”的转变。<sup>[21]</sup>

作为联结主义的高阶版本,机器学习的原理是:对于某类任务 T 和性能度量 P,在 T 上以 P 衡量的性能随着经验 E 而自我完善。<sup>[22]</sup>简言之,机器学习是一种面向开放的数据环境提取随机模式的方法(extract patterns from data)。它以环境中现存的法律语料而非以系统预设的专家规则为依据,实现一种基于归纳推理的反馈式计算。这是一种建立在学习性、认知性和统计学基础上的定量方法,其核心机理在于将开放获取的法律数据和之前形成的法律模式识别不停进行拟合,从动态的随机数据中建立临时的法律相关性模式,进行试错性的节点连接权重调整,从而不断逼近系统设定的法律理想值。<sup>[23]</sup>

因此,与符号主义相比,联结主义并不预设先天的法律专家知识规则,而是依靠后天的实际输出值与预设理想值的比对进行认知性调整。换言之,联结主义的心智观具有统计学特征,因此无需在法律因果关系层面对法律规范进行深度意义解释,而主要通过类似条件概率函数的方式进行预测性的计算。基于大数据的法律推理模型,即通过计算不同法律后果相对于法律构成要件的概率函数值,在一系列备选法律后果中找到概率函数值最高的一个,作为机器输出的法律答案。与符号主义的“外在语义论”相比,联结主义不需要设定法律世界的语义模型,因而省却了构造外部法律世界模型的麻烦;而与“莱布尼茨—乔姆斯基”理想语言模式相比,联结主义不需要刻画自然法律语言的深层语法结构,省却了对法律规范深层结构的分析。

更为激进的联结主义方案,则以罗德尼·布鲁克斯(Rodney Brooks)的具身化认知模型为代表。即认为智能系统应当完全对环境数据开放,甚至将环境本身作为系统决策的向导,由此,系统的内部符号表征不再重要,因为“世界就是认知系统所能有的最好的模型”。<sup>[24]</sup>这种理念意在建构一种无论在运作和认知上都对环境数据完全开放的系统,即认为在系统与环境之间存在直接的输入—输出关系,法律系统运作无需规范层面的封闭,在环境输入和系统输出之间,不需要任何符号层面的中介和转化,不需要任何逻辑规则的表征,也不需要内置的世界模型搭建。而这样一种对环境信息完全开放,直接面向世界的行为主义模型,实际已经预示了

[20] See Andrew Guthrie Ferguson, “Big Data and Predictive Reasonable Suspicion,” *University of Pennsylvania Law Review*, Vol. 163, No. 2, 2015, pp. 388-397.

[21] 有关大数据技术与法律预测, see Caryn Devins et al., “The Law and Big Data,” *Cornell Journal of Law and Public Policy*, Vol. 27, No. 2, 2017, pp. 388-401.

[22] 参见(美)米歇尔:《机器学习》,曾华军等译,机械工业出版社2008年版,第3页。

[23] 参见余成峰:“法律的‘死亡’:人工智能时代的法律功能危机”,《华东政法大学学报》2018年第2期,第8页。

[24] Rodney Brooks, “Elephants don’t Play Chess,” *Robotics and Autonomous Systems*, Vol. 6, No. 1, 1990, p. 5.

大数据主义法律智能的发展趋势。

## 2. 联结主义的主要问题

首先,联结主义(深度学习/大数据主义)本质上是一种统计学模型,其信息处理过程是对数据相关关系而不是因果关系的解释,因而不具备对法律规范语义的深层解释能力,无法在符号层面表征人类的法律推理过程。更确切地说,机器学习的“决策规则”是来自动态数据集的大量加权变量的概率函数,与人类法律推理和语义解释的规范风格存在内在冲突。<sup>[25]</sup>

其次,机器学习是以“化整为零”的方式,将宏观层面的法律解释任务分解为微观层面的数据并行加工。这种法律智能至多能获得现存法律知识图谱的一副统计学模拟图,而不具备反向修正法律知识图谱的能力,同时也缺乏根据当下遭遇的新问题迅速勾勒新的法律知识图谱的能力。只要相关的法律变化缺乏历史训练数据的支持,就不会得到系统算法的识别。完整的法律命题表征,被分解为大量的数据并行运算过程,因此,既没有法律符号表征的线性论证路线,也缺乏对于相关法律信息在系统内部处理过程的全局掌握能力,这自然会引发系统算法的黑箱性和可解释性问题。<sup>[26]</sup>

第三,表面来看,联结主义标榜更高的灵活性,彰显认知开放的特征,“自己不说话,让数据说话”。而实际上,联结主义采取的是数据“题海战术”,通过海量法律语料的训练,实现特定输入与输出的相关性模式识别。法律大数据系统在完成训练之后,一旦反馈学习达到收敛状态,就会迅速走向认知上的封闭,无法识别新类型的信息,除非启动新一轮的机器学习。深度学习一旦收敛,就可能出现“过拟合”现象,从而丧失认知上的灵活性。<sup>[27]</sup>换言之,法律机器学习的技术模型难以根据外部数据更新,自主调整系统节点的连结权重。对于新的法律问题,原有的系统难以处理,就只能将系统知识归零,重新设定理想值,针对更大规模的法律语料库,重新开启训练过程。

第四,在这种状态下,大数据系统缺乏运作的封闭性,只能被动跟随外部世界的的数据投喂周期,不断重启训练进程,以实现新的学习收敛。换言之,在联结主义模式下,智能法律系统既缺乏表征自身逻辑结构的元知识,也缺乏自我观察与自我纠错的能力。只能通过人类程序员来设定系统的理想值,依靠持续的法律数据投喂,依赖人工的深度参与和干预,借助强大算力和海量数据,以掩饰系统被不断归零重启训练的实质。正因如此,对于法律大数据的人工标注和智能系统“标准答案”的人为设定,又必然会引发算法偏见等问题。

[25] 参见(德)托马斯·维施迈尔:“人工智能系统的规制”,马可译,载泮伟江、翟志勇主编:《北航法学》(第8卷),中国政法大学出版社2023年版,第127页。

[26] 法律人工智能的黑箱性和可解释性问题, see Ashley Deeks, “The Judicial Demand for Explainable Artificial Intelligence,” *Columbia Law Review*, Vol. 119, No. 7, 2019, pp. 1829-1850.

[27] 法律人工智能的“过拟合”(overfitting)现象, see Frank Fagan and Saul Levmore, “The Impact of Artificial Intelligence on Rules, Standards, and Judicial Discretion,” *Southern California Law Review*, Vol. 93, No. 1, 2019, pp. 14-22.



## 二、法理学的古老命题：规范封闭与认知开放

符号主义范式,是从事先编制入系统的一般性规则知识出发进行封闭的运作;联结主义范式,是从海量开放数据出发寻找法律的相关性模式。两种人工智能法律范式,正呼应于法理学历史上实证主义法学与自然法学和社会法学的长期争论。实证主义法学坚持法律规范封闭的特征,自然法学和社会法学强调法律认知开放的维度。在光谱的绝对封闭一端,是纯粹以法律渊源为基础的推理,“法律就是法律”,法律系统之外与其相反的所有实质理由一概不进行考虑,以实现法律系统的运作封闭;在光谱的另一极端,则毫无例外地诉诸各种道德依据、原则和政策,以实现法律系统的认知开放。极端的规范封闭立场,可能导致机械适用规则(机械主义法学);极端的认知开放立场,可能导致法律适用缺乏确定性和可预测性(现实主义法学)。正因如此,任何法律系统都需要在最形式化的法律推理和最实质化的法律推理之间,找到一个最佳的制度平衡点。<sup>[28]</sup>对于法律人工智能,它同样需要综合形式性与实质性推理,打破实证主义法学与自然法学、社会法学的壁垒,在规范封闭与认知开放之间寻找到一条兼容二者的工程学路径。

遗憾的是,法律人工智能的既有范式延续了传统法理学在规范封闭与认知开放问题上二元对立的格局。符号主义亲和于形式主义法学,缺乏认知开放性;联结主义亲和于实质主义法学,缺乏规范封闭性。申言之,当前的法律人工智能陷入在规范封闭(符号主义)与认知开放(联结主义)之间各执一端的困境中,其共同症结则在于偏离了对“自创生”(autopoiesis)法律系统真实运作逻辑的理解,无法超越形式主义法学与实质主义法学的局限,融通不同法理学阵营的思想洞见。

首先,两种法律人工智能都预设系统运作所需的信息是充分的(无论是公理化的法律专家知识,还是人工标注的法律大数据),而这种“全能理性”的系统假设无法成立。在这种思想指导下发展的系统就会表现为对于环境复杂性的无知,无法在不确定的环境中自主修正自己的法律知识图谱。无论是符号主义还是联结主义,在实际应用中因此都沦为一个“无时间”(无变化)的机械系统。譬如,在符号主义范式下,所有的法律结论都被视为可以通过逻辑手段而一步步演绎出来的自然结果,这一法律结果甚至在演绎发生之前,实际就已经自动蕴含于规则前提之中,因此无法根据环境变化对这些前提进行实质性的修正与调整。而在联结主义范式下,一旦法律模型被训练完成,除了按照规定的模式对新输入的法律数据进行映射之外,这一系统的技术性能也不会发生根本变化。<sup>[29]</sup>

法律社会学的研究早已揭示,法律系统的规范封闭绝不等同于系统与环境的隔绝状态,法

[28] 参见(美)P. S. 阿蒂亚、R. S. 萨默斯:《英美法中的形式与实质》,金敏等译,中国政法大学出版社2005年版,第61页。

[29] 参见徐英瑾:“黑格尔‘逻辑学’对人工智能的启示”,《复旦学报(社会科学版)》2018年第6期,第36—37页。

律系统的边界是可以渗透的。法律系统需要在其内部结构中植入“可变性”要素,通过类似原则/例外这样的区分通道,灵活应对环境的变化。质言之,法律并非一台僵化的平庸机器(trivial machine),一台总是以相同且不断反复的方式,将输入转为输出的机器。相反,法律系统是一部历史的机器,法律系统每个自我再制的运作,都会改变系统的结构,将这部机器推移到另一种运作状态,并因此为进一步运作创造出已经变更的起始条件,“它会将自身的状态拉入到每个运作之中,透过每个运作而建构出一台新的机器”。〔30〕

与此相反,符号主义采取了将法律知识公理化的进路,因此就无法支持系统对于先天知识的自主性修正;而联结主义的人工智能一旦训练完成,也无法产生堪比人类法律智能的迁移能力。换言之,符号主义是以表面的规范封闭来排斥认知上的开放,联结主义则以表面的认知开放来遮蔽规范上的不封闭。而之所以说是“表面”,则是因为对于真实世界的法律系统来说,规范封闭与认知开放无法各执一端而存在。认知上的封闭,会最终瓦解系统运作的自主性(形式主义法学);规范封闭性的缺乏,则会导致认知开放丧失其定位的方向(实质主义法学)。因此,理想的法律人工智能,既需要规范封闭的演绎能力,也需要认知开放的学习能力。社会一方面需要智能法律具备形式理性,确保裁判的可控制性(规范封闭);另一方面也需要智能法律具备实质理性,以应对高度复杂的社会环境(认知开放)。智能法律必须同时担保运作上的封闭与认知上的开放,协调两者的内在张力。

从法理学视角观察,正是法教义学承担了化解法律的形式性/普遍性(规范封闭)与实质性/可变性(认知开放)二元张力的功能。换言之,对法律系统来说,认知上的开放恰恰必须以规范上的封闭作为前提。正是规范封闭的法律教义体系在决定法律系统如何进行认知的开放;决定法律系统如何进行资源和注意力的分配;决定相关事实和证据的认定;决定对它们的法律权重的配置。正因如此,实证主义法学才会认为,在司法过程中,即便存在指向相反结果的实质性依据,在该实质性依据成为制定法或先例之前,也必须仍然适用形式性规则。〔31〕环境信息必须通过法律规范网络的过滤,经过法律符号表征中介的处理,才能允许进入法律系统的沟通过程。

因此,只有规范上的封闭,才能形成稳定的法律语义推理结构,为法律系统的沟通运作确定逻辑边界,以对案件进行稳定有效的处理;只有规范上的封闭,外部的环境事件才能被转译为法律系统可以处理的信息;只有规范上的封闭,智能法律才能够化约环境复杂性,避免过度的认知开放,以免陷入“西西弗斯式的循环,不断给出新的语义推理结构,又在不断涌现的新证据的刺激下,不停推倒刚刚建立的结构”。〔32〕法律系统只对已在系统中得到语义界定的输入形式敏感,从而与环境保持距离,避免与外部世界的短路相接。由此,法律系统的规范封闭才可以担保实证主义法学的理想:规则越是形式化,才能严格限制司法专断的可能,从而保证同

〔30〕 (德)尼可拉斯·鲁曼:《社会中的法》,李君韬译,五南图书出版股份有限公司2009年版,第80页。

〔31〕 参见阿蒂亚等,见前注〔28〕,第9页。

〔32〕 徐英瑾:《心智、语言和机器——维特根斯坦哲学和人工智能科学的对话》,人民出版社2013年版,第241页。

案同判,创设法治事业的稳定性和可预期性。

申言之,智能法律不能对环境直接开放,否则将丧失系统的运作自主性。对于法律系统来说,所要处理的环境信息越多,反而越发依赖系统的规范封闭能力。系统只有在递归循环的自我指涉基础上,才能抵挡环境信息的过度激扰,获得对环境事件进行区分和标识的精致度。进言之,法律系统的规范封闭不仅可以排除环境噪音的干扰,它也借着使系统的敏感性集中于特定区域与议题,从而反向提升认知开放的能力。法律系统对于环境复杂性的隔离,恰恰有助于系统自身复杂性的提升,从而增强系统对环境的认知学习能力。

尽管深度学习技术擅长发现法律概念之间的相关性,但它无法理解法律概念之间真正的语义逻辑关系。大数据的概率统计,无法对输入和预测结果提供法教义学层面的解释。正因如此,研究者尝试通过符号方法的介入将统计关联转化为逻辑关联,从而增强因果关联的可解释性,进而形成法律人工智能的“神经—符号主义”(neural-symbolism)。<sup>[33]</sup>但这种尝试调和符号主义与联结主义各自优点的技术路线仍然停留在探索阶段,难以在基础理论层面回应两种范式的内在冲突。对于复杂法律系统的集成显然需要某种“顶层设计”,而当前的法律人工智能产业和实务应用,由于在符号主义范式遭遇瓶颈,因此更多是围绕法律大数据的采集和挖掘,在大数据主义的技术路径上狂飙突进。法律人工智能的认知开放有余,而规范封闭不足。

规范封闭无认知开放则空,认知开放无规范封闭则盲。法律人工智能遭遇的困境,其实正是法理学古老难题的延伸,即规范封闭/认知开放、形式推理/实质推理的系列二元张力。而当前的符号主义和联结主义模式,乃至两种技术路径的简单集成,都无法有效应对这种二元困境。这是当前人工智能法律面临的最根本性的挑战。正是两种技术路径在原理方面的不兼容,根本性地决定了简单集成无法解决真正的融合问题。这两条技术进路如何能够被有效衔接在一起?尤其是,法律人工智能目前主要从大数据主义的路径单兵突进(包括时下的法律大语言模型),这有可能导致在处理规范封闭和认知开放关系上的进一步失衡,颠覆法治作为规范性事业的价值,导向一种去规范化的技术治理模式,从而引发法律的深刻变异。我们可以从社会、时间和事实三个维度,考察由大数据智能技术引发的过度认知开放,有可能给法律发展带来的严峻挑战。

### 三、大数据智能技术驱动的法律变异

#### (一)智能法律个性化:社会维度的法律变异

大数据时代,已经催生了个性化医疗、个性化教育、个性化金融等现象,个性化法律似乎也很难豁免。法律正义的传统要求是法律面前人人平等,法律规则应当无差别地适用于所有主体。但事实上,正义的完整表达式是同等情况同等对待与不同情况不同对待。由于传统法律技术无法深入识别同等与不同的差异,难以精确定位不同的差异在法律上的相关性,因而也就难以针对不同情况在法律上做出不同的对待。为了避免法律上的“歧视”,传统法律主要采用

[33] 参见魏斌:“论新一代法律智能系统的融合性道路”,《法学论坛》2023年第3期,第22—26页。

统一化/普遍化的“算法”机制,有意保持一种认知的封闭,对法律对象采取同一化的形式推理模式。

但是,大数据智能技术的出现,则使法律的个性化和细粒化发展具有了可能,从一种传统上相对粗放的算法迈向一种更为精细的算法技术。传统法律进行统一的形式规则设定,通常不识别具体的个体差异,主要是立足对社会人口平均值的评估(例如受害人的平均伤害或加害者的平均预防成本),来设定类似于侵权法上“理性人”和“合理注意”的法律标准。这种对于平均值设定的法律偏好,其首要根源在于小数据时代信息获取的成本。

而当从小数据时代逐步迈向大数据时代,当数据采集和处理成本急剧下降,对于法律更为个性化和精确性的要求也就具有了可欲性和可行性。当法院和立法者可以更廉价迅捷地获取信息,也就为逼近更高程度的法律精确化,迈向一种“个性化”和“定制化”的法律做好了技术准备。从理论上来说,在大数据技术的加持下,制定个性化的法律规则,甚至可能比制定统一化的法律规则在成本/收益层面变得更具吸引力。<sup>[34]</sup>

申言之,在大数据智能技术的加持下,立法者不再需要精心设计一个要求聚合和权衡所有竞争性利益的单一形式规则。因为借助大数据技术,对于个人特征的数字画像,有可能会比对社会平均特征分布的统计更有根据和准确。<sup>[35]</sup>统一化的法律标准(例如对注意标准的设定)有赖于对社会平均值的计算,为了更准确地估算平均值,则需要利用统计学等技术,来全面地了解社会中非常复杂的技能和风险的分布状况。与此相反,如果只是针对特定的个人或临时标签群组成员,则可以更为便利地形成更精准的数字画像,制定“个性化”的注意标准,而不必在社会人口学的意义上,根据同类特征集合的平均分布,来设定“统一化”的法律注意标准。

在小数据时代,立法者和司法者无法采集足够数据来准确画像,因此只能依靠对平均值的估算来设定规则。例如高速公路时速 100 公里的限制,这一规则是基于对所有驾驶员的平均风险和技能的估计,以此来统一设定“理性人”的形式标准,并忽略任何个体的实质差异。小数据时代信息获取能力的欠缺,恰恰为形式主义的法律推理模式奠定了基础,认知能力的局限,限制了法律认知开放的程度。而在大数据时代,个性化法律则可以将每个司机切分为众多标签属性,一方面,可以同等对待每个属性(通过在算法中赋予其固定权重),另一方面,又可以累积计算针对某人的特定因素对待的积分总和,从而加总得出针对某人可以供给的个性化规则(例如个性化的时速限制或饮酒年龄标准)。<sup>[36]</sup>申言之,大数据时代的个性化法律将催生更实质化的法律推理技术,推动法律在认知层面采取更开放的态度,从而改变传统法律规范封闭的特征。

大数据智能可以深入识别个体的各种相关因素及其相对权重,由此根据立法者设定的目

[34] See Omri Ben-Shahar and Ariel Porat, *Personalized Law: Different Rules for Different People*, New York: Oxford University Press, 2021, p. 55.

[35] 参见余成峰:“数字时代隐私权的社会理论重构”,《中国法学》2023年第2期,第174—175页。

[36] See Ben-Shahar and Porat, *supra* note 34, p. 122.

标,将各种因素加总计算,从而生成个性化的规则。在这个过程中,基于大数据的智能法律推理将逐步进入不同法律场景,深入识别人际的相关差异,并根据这些实质差异,针对法律制裁、权利和义务作出个性化的规定。在大数据智能的视角下,传统法律算法当然是粗放的,因为统一的规则设定只能粗略地反映社会平均数,或只是多数派政治群体意志的体现,而大数据时代则为法律在社会面向的认知开放提供了新的可能和选项。

进一步的问题是,对于同等情况同等对待(同案同判)和不同情况不同对待(异案异判)这一正义公式的实现,大数据智能技术应当如何以前后一致的方式,并以公平的方式“对同等和不同做出区分”?什么事情应当被视为相同,什么事情应该被视为不相同?什么相同/不相同在法律上是相关的?判断的依据是基于大数据智能系统的自主设定(例如数字标签:相同标签相同对待/不同标签不同对待)?还是需要按照法律的规范标准设定?而无论是立足技术系统的认知性标准,还是立足法律系统的规范性标准,一旦形成个性化的法律对待,又会在何种层面构成新的法律歧视?这又将在多大程度上改变现代法律的规范封闭特征,侵蚀现代法治的形式理性标准?尤其是,这里还会涉及大数据的质量问题,一旦可用的输入数据越粗糙,越是接近识别某类特定群体的身份特征,那么智能系统就会将固有的社会歧视作为法律预测和决策的基础。更进一步,那些在大数据统计学趋势中被排除掉的异常数,往往很难通过重新调整训练集假设内容与样本空间尺寸的办法,在另一种概率表征的方式中得到表达。

质言之,由大数据智能技术推动的法律个性化,将在社会面向对法律统一性和普遍性的现代法治要求带来深刻冲击,挑战平等和民主这些启蒙价值,从而也将对法律的规范封闭/认知开放这一二元命题带来深远的影响。

## (二)智能法律预测化:时间维度的法律变异

大数据智能技术可能引发的第二种法律变异,更具颠覆性和破坏力。大数据技术的应用,不只会对社会面向对人际差异形成个性化的评估和对待,同时也会积极介入时间维度进行操作,形成法律的预测化。这种大数据的智能预测不同于传统的统计学建模,尽管两者都是为了应对未来的不确定性问题。因为,大数据法律预测具有独特的“述行性”的特征。<sup>[37]</sup>

古代社会运用占卜术来处理未来的不确定性问题(借助吉/凶卜卦对未来做出决断)。现代社会,主要是借助科学统计工具对未来进行计算。在统计学视角下,未来无法预测,但可以利用概率计算来管理和降低这种不确定性。而与占卜术和统计学相比,法律机制则是通过维护规范性期望稳定的方式,来应对未来的不确定性。即使法律的规范期望在未来面临失落,也仍要维持这一期望的稳定而不作认知上的调整。与之对比,大数据技术则尝试以预测未来的方式作为塑造未来的手段,从而迥异于现代统计学和法律的思路(对于两者,未来都是开放的、不可预测的)。申言之,大数据智能会预测规范性期望在未来发生失落的可能,由此提前改变这种规范性预期,或是通过先发制人的行动介入,积极干预这种预测在未来的发生概率。因

[37] 述行性(performativity),即奥斯汀(John Austin)所说的以言取效行为(the perlocutionary act),哈贝马斯称之为“通过在说某事时做某事而造成某事”。See Jürgen Habermas, *On the Pragmatics of Communication*, Cambridge: MIT press, 1998, p. 122.

此,基于这种对未来的预测,会进一步升级为对于未来的积极干预和预防,法律则会被作为干预和预防的治理工具,这将从根本上改变法律的传统功能和特征。

正如前述,传统的法律机制并不对未来进行预测,因此采取的是一种后发制人的时间视角:即只在“伤害”发生之后,才会通过制裁来干预个人的自由选择。而大数据智能技术则采取一种先发制人、防患于未然的时间视角,从而形成对特定时间、特定人群、特定主题和特定场景的普遍化预测,取向于对未来的风险观察、风险预防和风险排除。和现代统计学与现代法律不同,大数据智能不再只是处于被动等待的中立观察者和仲裁者,而会在时间面向积极介入、监控和干预世界。采用埃琳娜·埃斯波西托(Elena Esposito)的表述,大数据技术通过积极介入,主动地“制造”由自己预测的未来,预测由预测所塑造的未来。<sup>〔38〕</sup>并且,这种法律智能即使在预测上不准确,它仍然可以从实际的结果中学习,通过反向调整系统的参数和权重,以在下一轮预测中逼近目标。由此,大数据法律智能会同时采用①改变预测以适应世界;②改变世界以适应预测的双重策略,来减少预测失败的可能。<sup>〔39〕</sup>如此,法律的预测化就具有了认知性和规范性的双重特征,而这与法律的传统功能设定(维护规范性期望的稳定)完全背道而驰。

实证研究表明,依赖警务预测,最终会导致增加犯罪总量的悖论。<sup>〔40〕</sup>因为当大数据预测某地犯罪高风险,就会在此地增加警务投入,而其他地方则减少投入。高风险地区因为增加警务投入带来的犯罪率降低,可能仍然少于警务投入减少地区由此带来的犯罪率上升。更棘手的情况是,大数据的预测本身难以证实或证伪。如果该地犯罪率降低,有可能是本身犯罪风险并非如预测那么高;如果该地犯罪率进一步升高,则可以反向证明预测的准确性,因此进一步投入预防资源,或对预防方式进行调整。所以即使这种预测是准确的(后验),但很可能只是“自我实现的预言”。

这不仅会改变法律的规范封闭特征,甚至也可能构成对法律功能和意义的根本颠覆。因为法律不再是用来维护社会规范性期望的稳定,并以此保持未来的开放性,而是会被作为监控、塑造和支配未来的治理工具。时间面向的过度认知开放,将导致一系列预防性法律措施的出台,从而限制自由行动的可能性。问题不再只是错误预测的风险,而是由此减少了所有法律主体的未来行动空间。事实上,法律规则既不是对现实所做的描述,也不是对未来情况的预测,法律“并不是提供一个本属于现实世界的模式,而是要为现实世界提供一个模式”。<sup>〔41〕</sup>当人工智能成为一种预测未来的风险预防技术,就会从根本上侵蚀法律作为事后启动的规范性制度生存的空间。

〔38〕 参见(意)埃琳娜·埃斯波西托:《人工沟通与法:算法如何生产社会智能》,翁壮译,上海交通大学出版社2023年版,第116页。

〔39〕 同上注,第117页。

〔40〕 See Wim Hardyns and Anneleen Rummens, “Predictive Policing as a New Tool for Law Enforcement? Recent Developments and Challenges,” *European Journal on Criminal Policy and Research*, Vol. 24, No. 3, 2018, pp. 212-213.

〔41〕 (英)尼尔·麦考密克:《法律推理与法律理论》,姜峰译,法律出版社2005年版,第98页。

### (三) 智能法律风险化: 事实维度的法律变异

传统的法律制度必须在事实层面进行严格的界分,例如对特别极端的伤害来源和所有其他伤害来源严加区分,在法律推理中必须根据确定的证据阈值标准,确立严格的法律因果关系,不考虑一切与法律构成要件不相关的环境事实。只有当相关行为构成伤害的唯一或排他性来源,只有当行为比正常情况更有可能造成损失,才能满足法律的事实要件,承担相关损害的责任。无论是以财产边界纠纷为重点的物权法、以违背承诺为重点的合同法、以道德不法为重点的侵权法,都要在法律事实层面明确区分足以证明责任的异常行为和正常的免责行为。只有当侵犯财产红线,只有当不履行合同,或者侵害在道德上负有责任,相关的违约或损害赔偿才能成立。易言之,传统的法律制度在事实层面采取了一种全有全无/极化/二值化的规范封闭标准,从而大大限缩了法律在事实维度的认知开放程度。典型的法律事实必须是构成故意或者存在过失的行为,从而排除正常行为或仅有相关关系行为的法律责任。正因如此,一个更趋形式推理的法律体系较之更趋实质推理的法律体系,往往需要更有效的机制以查明事实真相。<sup>[42]</sup>

而在大数据技术的加持下,智能法律则可能在事实维度对损失原因或侵害来源采取一种截然不同的分析方法,从而深刻改变法律系统对于法律事实的认知态度,由此进一步加速晚近以来法律认知开放的发展态势。申言之,在风险社会的背景下,法律所关注的不再只是某种行动是否构成对他人的直接伤害,而是需要确保社会行动不会在任何意义上增加社会损失的风险。例如在现代合同诉讼中,问题已经完全围绕风险问题进行重新定位,法院面临的问题是哪一方应承担导致违约的条件变化的风险;而在侵权法上,可诉讼的问题是制造商是否适当地将产品伤害的风险降到最低。法院的作用不再只是惩罚违约行为,而是要在当事人之间分配潜在条件变化的风险;法院的关注点已经超出了特定的侵权动机,扩大到了更广泛的产品伤害风险。<sup>[43]</sup>

质言之,在法律风险化的观察视角下,社会所遭受的损失可以只是某种概率计算的结果。如果相关行动增加了发生损失的概率,就可以认定它们产生了损失。所有产生损失的行动因此可以被排列在一个统计学的连续体上,在概率为100%的损失(故意造成的损害)和概率为0%的损失之间,坐标轴上的所有分布样态都可能构成承担法律责任的概率相关性。“在风险贡献的连续体上,任何行动之间不存在明显的质的差异;所有行动都会带来风险。唯一的问题是贡献的程度。”<sup>[44]</sup>

在这种背景下,人工智能技术就会进一步催发法律成为一种风险控制工具。在实践中,大数据智能可以通过数字画像与风险精算,关注风险的每一个维度,对社会上所有以任何方式造成风险的因素进行数字监控,并推动法律作为内部化社会成本以控制风险的利器。由于每一项行动都会以某种方式增加损失的风险,法律的风险化因此也就意味着法律的规制范围将在

[42] 参见阿蒂亚等,见前注[28],第130—131页。

[43] See George Priest, “The New Legal Structure of Risk Control,” *Daedalus*, Vol. 119, No. 4, 1990, pp. 209-211.

[44] *Ibid.*, p. 215.

人工智能的辅助下获得极大的扩展,法律将更多以溯及既往、事前干预的方式呈现,以减少或分散任何意义上的社会风险。这同时也意味着,法律哲学将从规范封闭的义务论向认知开放的后果主义持续转变。

## 四、系统论视野下的法律人工智能新范式

### (一) 法律人工智能的三种理想类型:功能定位与类型区分

智能法律封闭性与开放性的矛盾,其实正是马克斯·韦伯(Maximilian Weber)著名的形式理性法和实质理性法之争的延伸。在韦伯看来,形式理性法与实质理性法存在永恒的冲突关系。形式理性法试图将所有透过分析而得的法命题加以整合,使之成为相互间逻辑清晰、不会自相矛盾、尤其是原则上没有漏洞的规则体系,所有可以想见的事实状况,全都合乎逻辑地涵摄于体系的某一规范,以免事实的秩序缺乏法律的保障。而实质理性法,则将伦理命令、功利的或其他目的取向的规则、政治准则等,作为法律决定的根本依据。在他看来,“法律逻辑的抽象的形式主义与他们欲以法律来充实实质主张的需求之间,(存在)无可避免的矛盾”。〔45〕

前文对智能技术过度认知开放催化的法律变异进行了说明,这也提示我们不能将人工智能法律的不同形态应用混为一谈,必须更准确地区分法律人工智能的不同功能定位与领域类型,对法律人工智能的多重理性逻辑做出更深入的分析,从而更好地协调规范封闭与认知开放的张力。笔者尝试借鉴韦伯的概念,在系统论视野下,构建人工智能形式法、人工智能实质法、人工智能反思法这三组理想类型,作为法律人工智能发展必须兼顾的三个方向。三类人工智能法应当各司其职,循环互动,共同构成具有内部分化与制衡关系的法律人工智能宪治结构的基础。

#### 1. 人工智能形式法/古典法

人工智能形式法是对法律人工智能最经典的想象。韦伯的法律自动贩卖机,对此做出了生动描绘:“人们从上头丢入事实(加上费用),它自下头吐出判决(及其理由)。”〔46〕其功能定位是,法律推理应该仅仅根据客观事实、明确的规则以及逻辑去决定一切为法律所要求的具体行为。〔47〕特别是在裁判简单案件(easy cases)时,一旦输入符合机器预置规则中构成要件的法律事实,经过规则知识库的指引,通过逻辑引擎推演,就能自动得出相应的法律结论。经由“输入事实—预设规则—推理结果”的线性处理模式,实现从法律事实向法律后果的自动转化。人工智能形式法的功能定位,因此乃是以运作封闭的形式推理,以法教义学的演绎逻辑,自动化地处理法律案件。借由法律系统的规范封闭,保障“形式化”的框架秩序,确立追求私人利益的自主行动领域,以主观权利的形式界分个人自由的活动空间。

但是,人工智能形式法需要面对的不只是简单案件,当遭遇困难案件(difficult cases 法律

〔45〕 (德)马克斯·韦伯:《法律社会学》,康乐、简惠美译,广西师范大学出版社2005年版,第220页。

〔46〕 同上注,第326页。

〔47〕 参见(美)史蒂文·J.伯顿:《法律和法律推理导论》,张志铭、解兴权译,中国政法大学出版社1998年版,第3页。



适用不明确)或者疑难案件(hard cases 适用结果不合理/不公正),此时,智能法律对于适用何种规则、如何适用规则以及适用规则的结果,都会面临如何选择以及选择本身的正当性问题。在多种备选规则、规则的多种解释以及适用结果的可接受/不可接受性这些问题上,面对不同的法律选项,就会不可避免涉及如何解释和决断的难题。此时,也就难以是人工智能形式法所期许的法律“唯一正解”的自动化推理。在这种情境下,智能法律的封闭运作,也就需要认知开放的奥援,以对困难案件和疑难案件做出合理的裁判。为了求得更高的合理性,人工智能形式法的可预测性,就可能需要让位于人工智能实质法的法律不确定性。

## 2. 人工智能实质法/规制法

对于人工智能形式法,最重要的是法律教义的融贯性和内在一致性,因此,并不关心智能法律是否适合社会规制的目标,而只考虑待审案件是否符合法律的构成要件,只需要根据“条件程式”(如果 A,那么 B)的设定,自动得出法律答案。其功能定位主要是事后的制裁和纠正。

而对于人工智能实质法,其目标首先不在于维持法律教义的内在一致性,而是政策导向和工具性的,只考虑在服务于特定规制目标时,智能法律的决策是否具有工具上的有效性。因此,人工智能实质法在认知上是高度开放的,具有宽泛的裁量权限,立足点不是维持运作上的封闭,而是追求政策目的和目标的最优化实现。主要通过收益/成本/风险计算以达成最优化目标,因此在功能定位上主要倾向于事前的预测预防、事中的监控干预,而不只是事后的矫正惩罚。

相比于人工智能形式法,人工智能实质法因此具有截然不同的法律分析方法,更接近规制法的含义。在人工智能法实质法的视野下,极端行为和正常行为的法律边界会变得模糊化,所有行动都可以被计算为对社会损失的连续贡献。法律因果关系的严格规范概念,因此会被相关关系的认知概念取代。人工智能实质法将主要致力社会议题化的问题解决,不会优先关注严格的因果关系和形式教义,而更多聚焦对实质问题的法律规制。

因此,人工智能实质法的规制领域,会从个人间的纠纷转变为司法社会工程学的场合。智能法律现在针对的不再只是特定当事人在特定时刻的异常行为,而是主要针对大数据意义上的特定类别行动者的代表,立足统计学的概率推断,计算相关主体是否处于更好的位置来预防风险或分散风险的成本。具体当事人之间的争议对于人工智能实质法是次要的,关键是如何最好地制定更广泛的规制措施以最大化社会福利。

## 3. 人工智能反思法

从规范封闭的角度来说,法律系统实现的是针对全社会的功能(function):即维护全社会规范性期望的稳定;从认知开放的角度来说,法律系统实现的是针对其他社会系统的功效(performance):即达到特定的社会规制目的。众多研究已揭示,在当代法律的社会功能和社会功效之间,已经出现愈益增大的张力,<sup>[48]</sup>这比韦伯在 20 世纪初叶对形式法实质化趋势的

[48] “通过给法律系统的内部能力设置限制,缓解法的功能和法的功效之间的紧张关系…法必须限制其对其他社会系统的功效。并非是在所有可能的社会情境中,在实质法的理性的意义上进行全面的纠纷调整。”(德)贡塔·托依布纳:《魔阵·剥削·异化:托依布纳法律社会学文集》,泮伟江、高鸿钧等译,清华大学出版社 2012 年版,第 304—305 页。

担忧要更为现实。更棘手的挑战在于,过度认知开放的人工智能规制法会产生追逐权力行为(power-seeking),寻求扩张自身对周围环境的战略影响力,这将带来难以预期的风险。<sup>[49]</sup>

在这种情况下,在人工智能形式法(类私法)和人工智能规制法(类行政法)之外,就迫切需要发展人工智能反思法这一具有宪治色彩的维度,以缓解法律功能/法律功效、规范封闭/认知开放、形式理性/实质理性的系列二元冲突。<sup>[50]</sup> 人工智能反思法,首先应当致力对人工智能规制法的元规制,分配与界定法律人工智能的控制权和决策权限,防止法律人工智能和社会的短路相接,确保人工智能追求与现代法治相匹配的目标,实现法律人工智能的价值对齐。其关键是保护数字基本权利,搭建可以容纳辩驳、商谈、异议和挑战权能的数字民主程序,激发法律人工智能的反思理性,通过具有自我反思宪治结构特征的设计,促进规制法“目的程式”的“再条件化”,制衡智能法律机器的黑箱化决策权力,限制智能法律规制功效的过度扩张,避免法律人工智能的过度实质化和去分化发展。<sup>[51]</sup> 如果说人工智能形式法是对应符号主义的正题,人工智能实质法是对应联结主义的反题,那么人工智能反思法就是规范封闭与认知开放二元兼容的合题。

## (二)封闭与开放的二元兼容:作为历史演化系统的法律人工智能

法律人工智能的两种范式如何兼容,并行不悖?判断人工智能法律的价值标准是其具有法教义学的融贯性(规范封闭),还是看其是否能够有效实现政策目标(认知开放)?法律人工智能应当通过形式逻辑进行推理,还是应当以实质的聚焦政策的方式进行推理?在前人工智能时代早已存在的形式理性法与实质理性法之争,在人工智能时代则会因为数字技术这一变量的深刻介入,呈现愈益复杂的面貌。

法律人工智能无法脱离法理学的深入研究而盲目发展,因为法律人工智能不得不预设某种有关法律之本质的看法,任何法律人工智能的技术开发,都必然对法律的性质和功能持有某种不自觉的信念。法理学与法律人工智能因此存在高度的相关性。任何对法律人工智能所做的技术说明,实际也在对法的性质和功能作出基本的假设;任何关于法的性质与功能的理论,也需要通过其在法律人工智能的工程实现和具体应用中进行检验。本文因此尝试从系统论法学的观察视域,从规范封闭与认知开放这一二元张力的兼容视角,超越传统法理学的二元对立格局,回应法理学在法律和社会问题上两极拉锯的难题,为人工智能法律的未来发展提供新的理论洞见和想象空间。

在系统论法学看来,法律并不是一个公理化的规则金字塔静态体系,而是具有高度动态运

[49] See Joseph Carlsmith, “Is Power-Seeking AI an Existential Risk?” 2022, pp. 1-57, <https://arxiv.org/abs/2206.13353>, last visited on 7 February 2024.

[50] 反思法的作用是通过人工智能法律系统的能力施加内在限制,调和功能和功效之间的固有紧张,设定法律人工智能的自我限制。“任何宪法的核心问题,都在于指涉自我、他者与全社会之间找到平衡。”(德)贡塔·托依布纳:《宪法的碎片:全球社会宪治》,陆宇峰译,中央编译局2016年版,第84页。

[51] 王锡锌以“数治”与“法治”二元框架,深刻揭示了数治权力对现代法治原则和价值的颠覆,即行政权过度扩张与侵害个人权益的双重威胁,并给出了法治约束数治的路线图,参见王锡锌:“数治与法治:数字行政的法治约束”,《中国人民大学学报》2022年第6期,第17—34页。

作特征和历史演化特性的节点化网络系统。正是通过法律规范节点之间的复杂动态连接,法律概念和法律命题的意义通过其在沟通网络中的拓扑学地位得以确定。系统论法学因此采取了一种运作建构主义的观察视角,任何法律表征的含义都可被把握为它在法律系统沟通网络中的拓扑学位置,因而独立并区别于外部环境的信息。法律系统的持续运作会实时参与系统网络中各个法律规范节点的连接权重分配。正因如此,每个法律规范节点的真值分配,并不由高高在上的“基础规范”决定,而是取决于相关的规范节点与其他规范节点之间的局域权重关联。具有高权重的法律沟通信道会获得暂时的稳定化,法律系统通过持续运作产生的冗余性,提升了法律问题的处理效率(规范封闭/同案同判),直至新的法律变异(认知开放/异案异判)推动法律系统的进一步演化,进而通过法律系统的自主选择机制,实现法律系统的再稳定。

质言之,自创生的法律系统既不是符号主义的逻辑系统,也不是联结主义的数据系统,而是历史主义的演化系统。法律系统唯有在实际的运作迭代中才能形成自主性和反思性。在这个意义上,法律系统应当被看成是一个动态循环的法律元素的再生产过程,这些再生产镶嵌在法律结构(规范)的超循环关系中。<sup>[52]</sup> 正因如此,对于法律人工智能,就不存在由人为的外部结构决定的可能性(无论是基于规则的专家系统,还是基于数据的深度学习)。真正智能化的法律系统,其法律结构的生成必然具有自创生的特征,诸法律运作为了能够递归地自我指涉其他运作,使系统自身获得不断界定,就必须预置高度抽象化的法律符号,将纷争事实涵摄为具有普遍性的法律类型,透过系统的持续运作来使法律沟通符号获得凝练与确认。进言之,法律系统是在其固有的历史状态中开启自身的诸多运作,法律因此不能被视为是对一个被预设为无漏洞的先天规则体系的适用,“法律沟通使用着系统的诸多结构,并且转变了这些结构的意义,以便能够在下一个运作中,重新使用这个先前已经被转化的结构,也重新再转化它”。<sup>[53]</sup>

在这种新的法理学视角下,法律就不是一个公理化的逻辑演绎系统。因为任何法律系统的知识和资源都是有限的,都需要在具体的法律问题求解语境中,根据当下裁判的需要进行实时的知识表征。法律概念绝非概念法学所理解的那种静态特征,而是在不断吸收外部环境信息的递归决策链条中动态更新其内容,每一个法律概念因此都嵌入了法律系统自我生产过程中的海量历史信息。<sup>[54]</sup> 这个过程必然是动态的,法律系统的语义表征会伴随沟通过程而发生历时的演变;这个过程也需要是充分灵活的,当新的法律事件和旧的规范网络不兼容,法律系统就需要恰当地修正原有的设定,通过“概念的扩张”与“解释的自由”,以适应新的外部经验。一旦封闭运作的法律符号系统彻底切断认知的开放通道,也就无法对环境变化作出自主的回应。也正是由于环境的高度复杂性,法律系统的知识库才必须伴随系统的历时性运作而不断修正和丰富。整个法律语义网的拓扑学结构,有可能会在任何一个时刻得到来自经验的修正,整个法律系统的结构因此处于动态的演变过程。正因如此,智能法律系统的具体运作形

[52] 参见(德)贡塔·托依布纳:《法律:一个自创生系统》,张骥译,北京大学出版社2004年版,第39—60页。

[53] 鲁曼,见前注[30],第73页。

[54] 参见宾凯:“从决策的观点看司法裁判活动”,《清华法学》2011年第6期,第104页。

态,一定是与该系统的实际运作历史紧密相关,而不由该系统搭建时内置的一般推理规则完全决定。

搭建具有自主演化和自我反思能力的人工智能法律系统,实现规范封闭与认知开放的二元兼容,也需要建立包含智能司法、智能立法、智能规制与智能合约的多重维度,以模拟真实法律世界以司法系统为中心,以立法、规制和合同机制为外缘的结构。<sup>[55]</sup>以智能司法作为智能法律系统的中心,实现规范上的封闭;以智能立法、智能规制与智能合约作为智能法律系统的外缘,实现认知上的开放。法律系统的外缘适合于作为与全社会其他功能系统的接触地带,在这个缓冲带,各种环境激扰可以被带入到法律形式之中。通过这种结构耦合的中介装置,一方面可以限制环境条件对法律系统的直接影响,另一方面则可以提升法律系统与外部环境的共振能力。通过智能立法、智能规制与智能合约的边缘接触带,智能法律系统可以为来自环境的信息建立一个高度敏感的接收站与转化站。智能法律系统的规范封闭,因此可以避免环境对系统的直接输入(基于大数据的法律推理模型,包括时下热门的大语言法律模型,均是此类版本的变种),而是经由智能法律系统与其他社会系统的结构耦合形成认知开放能力。一方面,以中央化的智能司法系统为中心,维护规范性期望的稳定,避免系统直接回应环境;另一方面,建立人工智能法律系统与政治系统(立法/行政)、经济系统(合同)以及心理系统(主观权利)的结构耦合机制,接受来自环境和其他社会系统的外部激扰,保持认知上的敏感性和开放性。

搭建具有自主演化和自我反思能力的人工智能法律系统,实现规范封闭与认知开放的二元兼容,也需要处理组织化领域与自发领域的分化问题。<sup>[56]</sup>其目的是防止法律人工智能系统被任何单一的中心支配与控制,助推智能法律系统形成分布式的反思性商谈结构。其关键在于促成法律人工智能系统的内部分化,确保组织化、职业化部分的法律人工智能无法对整个系统形成全面支配。各个专业领域以中心化的方式决断,但同时这些领域又经由各种外部压力和自我的反思学习,将其他领域的规范以及整体秩序的规范,尤其是将来自人类公共领域的沟通反馈,吸纳进人工智能法律沟通的内部视角。

在这种新的法理学观察视角下,法律人工智能就需要实现一种以司法为中心的普通法运作模式,个别案件必须在一个由先前与稍后的诸法律决定所构成的二阶观察脉络中,使自身获得定位。智能系统需要建立一种基于案例的类比式推理模型,形成具有内部复杂性的由支撑论据之权衡所构成的递归性沟通网络,这个系统应当具有高度弹性的反思概念装置,透过系统的案件区分与论证技术,而非以公理体系化的方式进行抽象的演绎推理。这一智能法律系统,可以对过去的经验进行学习,并能够在资源约束的条件下对给定的问题进行实时解答。任何一个法律概念的语义都置于概念相互关联的沟通网络之中,伴随法律系统运作经验的不断丰

[55] 卢曼认为,现代法律的生产方式有三种:立法、司法和订立合同。司法之所以处于中心,是因为只有司法才有义务(被强迫)在信息不全、充满不确定性的条件下仍然必须做出裁决。参见上注,第97—98页。

[56] 组织化领域与自发领域的分化,在经济系统体现为企业/消费者,在政治系统体现为政府/公共意见,在法律系统体现为司法/被法律约束的人,有关社会系统的宪治领域分化问题,可参见托依布纳,见前注[50],第104—113页。

富而被修正,进而发生系统拓扑学结构的变化。在规范和事实都具有可变性的情况下,一方面,规范向事实下延,另一方面,事实向规范提升,实现规范与事实的相互内建。规范封闭与认知开放因此不再是二元对立的关系,而是伴随系统的运作演化而获得各自能力的同步提升。在这种状态下,当法律人工智能面对“困难案件”,也就是当环境事实使得系统旧有的规范网络无法直接取舍,因此就必须动员更大范围的规范网络节点;而当面对“疑难案件”,则会激发法律系统着手修正原有规范网络的拓扑学结构。这两类案件,都会催发中央化的智能司法系统以规范封闭的方式,来回应由认知开放所带来的系统不融贯性问题,进而提升自我反思和自我创生的能力。

与此同时,也需要持续部署分布式的人工智能法律系统,在这些不同的智能法律系统之间,形成网络化的法律论证的竞争、沟通与学习机制,进而建立可以处理不同系统之间法律冲突的准据法程序,建构系统内部的反思概念装置及其外部协调的审查整合功能,以模拟现实法律世界的司法层级化组织,由此解决智能司法决策的既判力和正当性问题。

---

**Abstract:** Symbolism and connectionism are two classic paradigms of artificial intelligence, leading to rule-based and data-based legal reasoning models, respectively. These models correspond to the age-old jurisprudential proposition of normative closure and cognitive openness. Symbolism emphasizes the closure of norms, while connectionism focuses on the openness of cognition. These two paths represent distinct technological paradigms, underpinned by deep philosophical differences, leading to two entirely different legal development models. Thus, the openness and closure of legal AI encompass a range of important dualistic jurisprudential propositions. Around this dual concept, within the intellectual traditions of formal rational law and substantive rational law, using the conceptual tools of systems theory jurisprudence, the core theoretical propositions of AI in law can be reconstructed. The unilateral drive of big data intelligence technology will provoke profound legal mutations across the social, temporal, and factual dimensions of law. This demands a reevaluation of the functions, domains, processes, and structures of legal AI, defining three types of evolutionary paths. From the dual perspective of normative closure and cognitive openness, more complex legal theories should be developed to design the developmental roadmap for legal AI, guiding the healthy evolution of intelligent law in the digital age.

**Key Words:** Legal Artificial Intelligence; Normative Closure; Cognitive Openness; Personalized Law; Reflective Law

---

(责任编辑:王锡铤)