

基于裁判文书数据的法实证研究方法

——以“知假买假”问题为例

熊丙万 王军乐

摘要：以“知假买假”问题为例，通过对惩罚性赔偿裁判文书的信息抽取和数据分析，评述了基于裁判文书数据的法实证研究方法。法实证研究是一个完整的循环过程，它以法学的规范性问题意识为起点，在目的论或结果论导向下，结合问题意识、理论资源、实证材料对问题与概念进行拆解，构建起连接法学问题意识与实证信息的桥梁，进而通过实证材料提升事实判断的品质，从而更好地回答相关法学规范问题。裁判文书是重要的实证材料，利用大规模裁判文书进行实证研究时，可以通过机器学习等技术手段从裁判文书中准确地抽取信息。同时，裁判文书所含内容是法律实践活动的重要但有限的体现。基于裁判文书进行实证研究时应当谨慎地将其用于数据分析，严格评估文书数据质量、灵活采取替代性分析方案。与其他科学研究一样，法实证研究不提供对经验世界的确定性结论，而是追求盖然性相对更高的解释，并在开放的态度中不断发展。

关键词：实证研究；裁判文书；信息抽取；惩罚性赔偿；知假买假

[中图分类号] D922.294 [文献标识码] A [文章编号] 2096-6180 (2023) 06-0116-22

引言

对法律问题的实证分析 (empirical analysis of legal issues) 有三个分支，运用实证方法描述法律制度在经验世界的运作情况 (即“法实证研究”⁽¹⁾) 是其中之一。⁽²⁾ 法实证研究旨在探究与“法”有关的各种事实，描述“法”在现实世界运作的过程、效果等。法实证研究受到了自然科学实证

【作者简介】熊丙万，法学博士，中国人民大学法学院副教授；王军乐，中国人民大学未来法治研究院研究人员。

【基金项目】国家社会科学基金一般项目“职业打假的社会效应与司法政策研究” (项目批准号：21BFX109)。

(1) 此类研究过去常称为“法律实证研究”。关于“法律实证研究”这一称谓的争议与批评参见侯猛：《实证“包装”法学？——法律的实证研究在中国》，载《中国法律评论》2020年第4期，第61-71页；关于“法实证研究”这一称谓的合理性讨论参见彭小龙：《法实证研究中的“理论”问题》，载《法制与社会发展》2022年第4期，第201-224页。

(2) 另外两个分支为诉讼当事人利用自然科学实证方法 (如DNA鉴定)、在个案中运用社会科学实证方法 (如证明是否存在种族歧视)。Theodore Eisenberg, *Empirical Methods and the Law*, 450 *Journal of the American Statistical Association* 665, 665(2000).

研究方法的启发但并非简单套用。在法实证研究中,实质命题(法学的规范问题)与方法问题、观察与理解相互交织,本质上仍属于法学理论研究。⁽³⁾就法实证研究与规范研究的区分而言,法实证研究的核心是数据等经验材料。美国部分法律学术同行采用数据定量分析的 ELS (Empirical Legal Studies) 逐渐将法实证研究推向了一个新的层次。

裁判文书的上网公开为观察我国的司法运行提供了一扇新的窗口。虽然一股遍布国内外的关于中国法院的数字学术浪潮已经开始⁽⁴⁾,但是目前的一些基于裁判文书数据的法实证研究成果在领域广度⁽⁵⁾、数据规模⁽⁶⁾、信息挖掘深度⁽⁷⁾、数据可靠性披露⁽⁸⁾等方面仍存在不足。

网上公开的裁判文书文本仅是裁判文书的数字化形式,未实现裁判文书信息的数据化。在分析裁判文书数据前,需要先从每一份裁判文书中抽取信息并转化为符合分析需要的结构化数据。对海量裁判文书进行信息抽取离不开计算机技术的辅助。⁽⁹⁾目前,国内外学界对裁判文书信息抽取和数据运用都已有讨论⁽¹⁰⁾,但尚无文献面向法学研究者提出一套系统化的操作方法。

本文以法学中的“知假买假”问题(详见后文“问题意识的确立”)为例,以消费者惩罚性赔偿全量裁判文书为样本,系统论述在问题意识的框架下面向法实证研究的裁判文书信息抽取流程与方法⁽¹¹⁾,并探讨运用裁判文书数据分析法学问题的方法。基于裁判文书数据的法实证研究的整体思路如图 1 所示。

(3) 彭小龙:《法实证研究中的“理论”问题》,载《法制与社会发展》2022年第4期,第223页。

(4) Benjamin L. Liebman et al., *Mass Digitization of Chinese Court Decisions: How to Use Text as Data in the Field of Chinese Law*, 8 *Journal of Law and Courts* 177, 197(2020).

(5) 目前,我国基于裁判文书数据的法实证研究所涉学科领域之中心在于刑事诉讼,其他领域的研究则极其缺乏。参见屈茂辉:《基于裁判文书的法学实证研究之审视》,载《现代法学》2020年第3期,第34-36页。

(6) 既有的基于裁判文书数据的研究一般仅包括几十份到几百份的裁判文书,例如葛江虹:《“知假买假”:基于功能主义的评价标准构建与实践应用》,载《法学家》2020年第1期,第160-172、196页。该文分析了产生于2017年的459份适用《消费者权益保护法》第55条第1款的裁判文书,发现在针对“假冒伪劣”或“不实陈述”问题和针对“不规范标识”问题的知假买假索赔中,法院更支持前者。但被归类为“假冒伪劣”和“不规范标识”的案件数量分别只有47件和49件。

(7) 案由、裁判结果、文书引用的法条等浅层信息往往无法满足论证需要,数据分析沦为装点陪衬。

(8) 不仅具体变量的编码与界定模糊,信息抽取的具体方法和结果评估亦语焉不详。对于信息抽取方法,一般研究仅说明运用了何种技术方案,如“自然语言处理模型辅之以人工校正”(参见马超、郑兆祐、何海波:《行政法院的中国试验——基于24万份判决书的研究》,载《清华法学》2021年第5期,第196页)、“基于无监督机器学习的主题模型(topic model)结合选择性的深入阅读”(Benjamin L. Liebman et al., *Mass Digitization of Chinese Court Decisions: How to Use Text as Data in the Field of Chinese Law*, 8 *Journal of Law and Courts* 177, 190-195(2020))等,或干脆不提技术方案(参见白建军:《论刑法教义学与实证研究》,载《法学研究》2021年第3期,第117-132页)。对于信息抽取结果,读者也难以获知其结果在多大程度上全面、准确。

(9) 信息抽取(Information Extraction, IE)是指按照预先设定的目标从文本(一般是非结构化的自然语言文本)中自动化地抽取特定的信息并形成结构化的数据。参见程显毅、朱倩、王进编:《中文信息抽取原理及应用》,科学出版社2010年版,第3-10页。

(10) 例如 Benjamin L. Liebman et al., *Mass Digitization of Chinese Court Decisions: How to Use Text as Data in the Field of Chinese Law*, 8 *Journal of Law and Courts* 177, 177-201(2020); 邓矜婷:《计算法学方法初阶》,法律出版社2021年版,第166-208页。

(11) 之所以强调“面向法实证研究”,是因为判决预测、类案检索、基于数据的智能裁判等领域的研究也需要对既有裁判文书进行信息抽取,虽然三者的信息抽取有共通之处,但侧重点和目标存在不同。

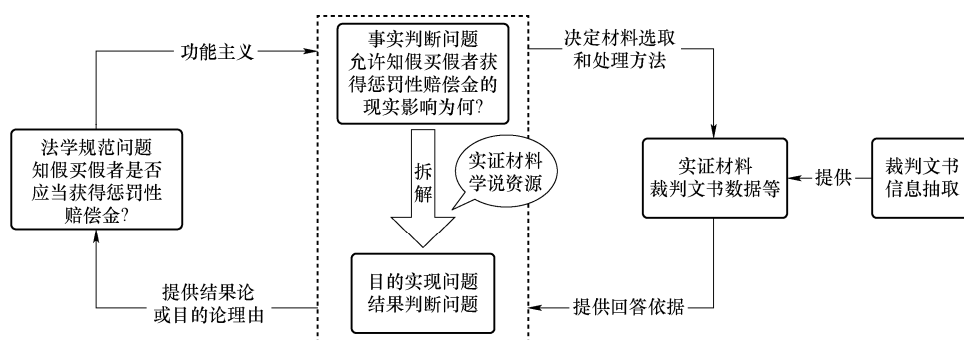


图1 基于裁判文书数据的法实证研究思路

下文首先论述研究的理论准备和材料准备，即在理论上分析系统的法实证研究的价值和必要性，并分析裁判文书数据作为法实证研究材料的优势、局限及应对。接下来讨论关于与知假买假相关的法学规范问题如何经由功能主义导向事实判断问题以及事实判断问题的拆解。之后着重讨论裁判文书的信息抽取方法。最后讨论通过裁判文书数据得出事实判断结论的过程中存在的问题和可能的解决方案。

一、研究的准备：法实证研究中的理论与材料

(一) 事实判断问题的规范价值

法实证研究者经常需要面对诸如“法实证研究是否必要”“实证研究在法学研究中有何作用”等问题，有必要对实证研究与规范研究之间的关系以及实证研究的价值作出解释。

法学研究中的事实判断问题是指有法学研究价值的经验问题，如现实世界中存在哪些利益冲突、对这些利益冲突过去采取的协调策略及其效果、某一具体规范的事实基础等。的确，法律规范的体系构建在任何时候都是至关重要的基础性工作，但其毕竟不是真空中的思想游戏，法律规范是嵌入现实生活中的。有同行甚至夸张地说：“如果没有对中国社会做很好了解就进行法律体系建设，法教义学的工作越是成熟、技艺越是精湛，法律实践的荒谬性就越强。”^{〔12〕}“是”不等于“应当”，实证研究没有也不应把“是”当成“应当”来处置。事实判断问题对于规范研究的价值主要体现在以下两个方面。

一是对某种社会关系的存在与演变这类问题的讨论结论，直接决定着后续规范讨论的必要性。如果不存在信息时代的个人信息收集、处理、利用事实以及由此产生的利益冲突，也就没有必要讨论个人信息保护规范。在一定意义上，民法典制定中所有的民法典之“变”，都以中国之问、时代之问、共识之变等事实判断之“变”作为前提和基础。^{〔13〕}

二是事实判断问题的结论经常为具体法律规范论证提供结果论或目的论的说明，即某规范会

〔12〕 程金华：《当代中国的法律实证研究》，载《中国法学》2015年第6期，第77页。

〔13〕 王轶：《民法典之“变”》，载《东方法学》2020年第4期，第41页。

导致什么样的结果,或某规范能否实现预定目的。^[14] 结果论或目的论的论证通常需要在因果关系层面作判断。因果关系中的“原因”可以指结果发生的必要条件、充分条件或结果发生过程中既不必要也不充分但很关键的因素。当在第三种意义上使用因果关系时,这种因果关系可以是概率性的、包含背景条件的。^[15] 例如,堕胎合法化解释了美国 20 世纪 90 年代犯罪率下降的 25%~30%。^[16] “堕胎合法化是犯罪率下降的原因”这个判断是概率性的,同时包含了“因不能合法堕胎而降生的孩子更有可能无法得到良好的照护和教育,而堕胎合法化避免了这些孩子的出生”等背景条件。由于“犯罪率的下降是应当追求的”,所以“在特定背景下,堕胎合法化会导致未来犯罪率的下降”这一事实判断结论为“应当允许堕胎”这一规范命题提供了一个结果论的理由。^[17]

针对“知假买假”问题,以事实判断结论作为政策主张的目的论或结果论的前提是常见的论证路径之一。有研究者以新闻报道为依据,试图以“现行惩罚性赔偿制度存在消极结果”(结果论)和“奖励举报制度有助于市场监督目的的实现”(目的论)论证“以奖励举报取代知假买假”的合理性,但其对这两个事实判断命题的回答依据薄弱。^[18] 若要把握现行制度下知假买假行为的整体特征,需要通过系统的实证研究得出更加可靠的结论。

(二) 系统的实证研究何以必要

虽然我们不能从被观察的事例中通过归纳推理得出确定性的因果关系,但观察的数量越大,我们观察到的关联越有可能反映真正的因果关系,结论的盖然性通常就越高。^[19] 因此,法实证研究不以个别的、预设的、理想中的模式为参照,而是通过收集和分析数据,从整体上观察、描述、预测法律现象。^[20] 宏观、复杂的法律现象往往超越了研究者能够直观把握的范围,难以通过简单的观察获得对法律现象及其因果关系的准确认知,因而需要通过系统的实证研究方法,广泛收集

[14] 当作为结果论的说明时,论证形式为:“采取规范 N 会导致某个结果 R,而 R(不)应当追求,所以(不)应当采取 N”;当作为目的论的论证时,论证形式为:“某目的 P 应当实现,而采取规范 N 有助于(有害于)实现 P,所以(不)应当采取 N。”“N 是否会导致 R”或“N 对实现 P 有何种作用”都属于事实判断问题。欲得出是否应当采取 N 的结论,需要先回答这些事实判断问题,事实判断问题的结论提供了采取或不采取某种立法论或解释论的一个理由。参见王鹏翔、张永健:《经验面向的规范意义——论实证研究在法学中的角色》,载《北航法律评论》2016 年卷,第 50-51 页。

[15] [美] 欧文·M·柯匹、[美] 卡尔·科恩:《逻辑学导论》(第 13 版),张建军等译,中国人民大学出版社 2014 年版,第 543-546 页。

[16] 作者通过系统的实证研究发现,20 世纪 90 年代美国犯罪率的下降不是源于经济的繁荣、人口结构的变化、更好的刑事政策、更严的枪支管控、枪支隐蔽携带法案、死刑的增加,而是源于警力的增加、监禁人数的增多、快克可卡因毒品的遏制、1973 年罗诉韦德案后的堕胎合法化。Steven Levitt, *Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not*, 18 *Journal of Economic Perspectives* 163, 163-190 (2004).

[17] 是否应当限制堕胎的问题还应当考虑其他多种因素,但已超出本文讨论范围,不再赘述。

[18] 作者所称的“奖励举报”即知假买假者发现问题商品或服务后向国家机关举报,国家机关查实后给予举报者物质奖励。作者引述了涉及 18 家企业的新闻报道,报道称这些企业近两年遭遇的职业打假索赔中涉及包装标签问题的索赔比例最大,达 55.78%,而涉及产品质量问题的仅占 11.79%。这些少量样本对于知假买假现象是否有足够的代表性存疑。作者主张的“奖励举报可以鼓励知假买假人分散‘打假’的范围,更多关注商品质量问题”则没有实证依据。参见刘保玉、魏振华:《“知假买假”的理论阐释与法律适用》,载《法学论坛》2017 年第 3 期,第 66 页。若真的以奖励举报取代知假买假制度,很可能刺激部分举报人广泛举报,浪费行政执法资源。

[19] [美] 欧文·M·柯匹、[美] 卡尔·科恩:《逻辑学导论》(第 13 版),张建军等译,中国人民大学出版社 2014 年版,第 579 页。

[20] 白建军:《论法律实证分析》,载《中国法学》2000 年第 4 期,第 34 页。

信息并以规范的过程进行科学分析，从而最大程度地提高结论的可靠性。⁽²¹⁾

曾经，我国的法实证研究对方法的重视程度并不高。⁽²²⁾但随着法律大数据逐渐进入法实证研究的视野，诸如如何建立规范命题与经验命题的关联、如何保障获取数据的可靠性、如何处理和分析海量的法律数据、如何避免“统计撒谎”、如何明晰数据解释的边界等问题越发尖锐地摆在实证研究者面前。⁽²³⁾实践中甚至出现了两位研究者针对同一问题的实证研究结论存在明显差别甚至相互反对的情况。⁽²⁴⁾可见，实证研究的可靠性不仅需要充足的数据，也需要科学的方法。

相比自然科学，社会科学中的普遍现象观察与因果关系判断更为困难。法的实际样态是法的观念价值、制度规范以及各种社会要素互动的结果，涉及行动主体在规范认知、内容、动员等多层面的复杂互动，而研究者也置身于法律参与构造的社会中。⁽²⁵⁾因此，研究法学中的事实判断问题更需要系统化的实证研究方法。尤其是对于法的实际样态的研究，只能通过实质命题与方法问题、观察与理解等交织而成的“分析框架”为人们所认识。⁽²⁶⁾

(三) 裁判文书数据在法实证研究中的优势

掌握经验（收集数据）是进行实证研究的重要前提。⁽²⁷⁾实证研究中的数据主要有四种来源：裁判文书、通过实验创造的经验、研究者自行收集的一手资料、官方公布的统计数据。⁽²⁸⁾在法实证研究中，法律实验目前难度较大且成本高⁽²⁹⁾，研究者能够收集到的主要是另外三类非实验数据。

通过访谈、观察等方式收集一手资料成本较高且范围有限，主要用于案例研究或定性研究。司法统计数据（国家机关工作报告、法律年鉴、统计年鉴等）则往往难以契合研究需要。⁽³⁰⁾

裁判文书数据有信息丰富、覆盖广泛、数量规模庞大和获取便利等优势。裁判文书是法院司法实践的直接产物，记载了诉讼参与人、程序、事实、论证、裁判等信息，可供多角度研究司法

(21) 系统的实证研究要求过程的规范性，只有研究的过程是规范、清晰、可重复的，其对事实的描述和解释才是可信的。作为一种研究方法，实证分析由三个基本要素构成：程序、经验、量化。参见白建军：《论法律实证分析》，载《中国法学》2000年第4期，第30页。

(22) 有学者认为实证研究的方法这类话题并不重要。参见苏力：《好的研究与实证研究》，载《法学》2013年第4期，第16-17页。

(23) 程金华：《迈向科学的法律实证研究》，载《清华法学》2018年第4期，第151页。

(24) 左卫民：《“热”与“冷”：非法证据排除规则适用的实证研究》，载《法商研究》2015年第3期，第151-160页；易延友：《非法证据排除规则的中国范式——基于1459个刑事案例的分析》，载《中国社会科学》2016年第1期，第140-162、206-207页。针对“启动非法证据排除程序的案件在申请非法证据排除案件中所占比例”这一问题，左卫民文得出的比例是48.1%，易延友文得出的比例则是89.37%。

(25) 彭小龙：《规范多元的法治协同：基于构成性视角的观察》，载《中国法学》2021年第5期，第167-172页。

(26) 彭小龙：《法实证研究中的“理论”问题》，载《法制与社会发展》2022年第4期，第221页。

(27) 在数据作为经验世界中信息的载体和表现形式的意义上看，除通常所说的数值形态的数据形式外，通过访谈等方式收集到的资料（个别经验）也是数据，只不过是而非结构化的、残缺的数据。

(28) 黄国昌：《法学实证研究方法初探》，载《月旦法学杂志》2009年第175期，第145-152页。

(29) 实验的关键在于控制变量，但法律实验以人为对象进行，其中的变量难以控制。更重要的是，若以虚构的场景进行实验则意味着缺乏现实的利益冲突关系，实验中的当事人缺乏对现实场景中当事人的代表性；若对现实中的利益冲突采取不同的对待方式，则有违平等的要求。实践中出现的个别有限的“法律实验”主要是通过各种手段观察现实的法律实践，而非直接创造经验。参见唐应茂、刘庄：《庭审直播是否影响公正审判？——基于西部某法院的实验研究》，载《清华法学》2021年第5期，第146-162页。

(30) 司法统计数据存在统计口径不统一、公开不够全面及时、数据不连贯、公开渠道零散等不足，无法形成相互贯通、可互相比对的数据网络。参见易霏霏、马超、何海波：《我国司法统计数据的公开：现状与建议》，载《中国应用法学》2017年第2期，第69页。

实践的实际运作。⁽³¹⁾截至2022年9月5日,中国裁判文书网(<https://wenshu.court.gov.cn/>)已公开135 704 431份裁判文书。除中国裁判文书网之外,市面上还存在多个裁判文书商业数据库。通过“爬虫”⁽³²⁾或与相关机构合作等方式,研究者可以相对容易地获取大量裁判文书。

(四) 裁判文书数据在法实证研究中的局限

虽然利用公开裁判文书数据进行法实证研究有着多方面的优势,但大数据不是万能的。科学研究需要认识到自己的局限性,为自己的命题澄清前提、划定界限。

首先,裁判文书形成时存在非随机性。发生纠纷后,除了由法院判决外,当事人还可以自行和解,法院也可以主持调解⁽³³⁾,因此通过司法审判解决的纠纷只能代表现实生活中纠纷的一部分。一项法经济学研究表明,法律纠纷中的双方当事人对于判决结果的分歧越大、纠纷事实越接近法律适用标准的边界(即存在更大的辩论和争议空间),该纠纷就越可能通过诉讼来解决。⁽³⁴⁾

其次,裁判文书公开时也存在非随机性。不同年度、不同案件类型与文书类型、不同法院层级和案件审级、不同地域的裁判文书的上网公开率都有所不同。⁽³⁵⁾关于造成裁判文书公开率差异的原因,尚无被广泛认可的解释。⁽³⁶⁾热点、敏感案件的裁判文书公开比例仍然明显偏低。⁽³⁷⁾此外,文书公开后也可能被撤回。司法解释关于裁判文书撤回的规定相对简略⁽³⁸⁾,实践中的裁判文书撤回主要有当事人申请撤回和法院主动撤回。绝大部分当事人均不同意自己案件的裁判文书上网公

[31] 通过分析诉讼当事人及其行为可以发现,当事人拥有的身份、财富等社会资源并不直接影响法院判决,但将这些资源投入诉讼中则会影响判决。参见田燕梅、徐恺岳、魏建:《法院判决的影响因素——当事人资源与诉讼能力的中介效应》,载《清华法学》2021年第5期,第163-176页。通过分析法官说理可以发现,法官适用民法时存在“分散解释”“另类解释”“异常解释”等非典型解释方案,影响民法典的统一实施。参见许可、张永健:《论民法典的统一实施——理论辩正与实证分析》,载《清华法学》2021年第5期,第89-110页。

[32] 关于通过爬虫技术获取裁判文书的方法,参见邓矜婷、张建悦:《计算法学:作为一种新的法学研究方法》,载《法学》2019年第4期,第113-114页。

[33] 根据《最高人民法院关于人民法院在互联网公布裁判文书的规定》(法释〔2016〕19号)第4条第3项的规定,除为保护国家利益、社会公共利益、他人合法权益确有必要公开的外,以调解方式结案或者确认人民调解协议效力的文书不在互联网公布。

[34] George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 *The Journal of Legal Studies* 1, 14-37 (1984). 有学者对诉讼的非随机性的存在提出了质疑,其根据是和解赔偿金和判决赔偿金在某些方面的统计相似性, Eric Helland, Daniel Klerman & Yoon-Ho A. Lee, *Maybe There Is No Bias in the Selection of Disputes for Litigation*, 174 *Journal of Institutional and Theoretical Economics* 143, 143-170 (2018). 但对这种相似性是否能够否定诉讼的非随机性,存在不同意见, Jonah B. Gelbach, *Maybe There Is No Bias in the Selection of Disputes for Litigation: Comment*, 174 *Journal of Institutional and Theoretical Economics* 171, 171-176 (2018).

[35] 杨金晶、覃慧、何海波:《裁判文书上网公开的中国实践——进展、问题与完善》,载《中国法律评论》2019年第6期,第128-133页。

[36] 关于造成裁判文书公开率差异的可能原因的讨论,参见 Benjamin L. Liebman et al., *Mass Digitization of Chinese Court Decisions: How to Use Text as Data in the Field of Chinese Law*, 8 *Journal of Law and Courts* 177, 185-190 (2020); 唐应茂:《领导意愿、机构能力和司法公开——北京、上海、广东裁判文书上网率的初步研究》,载《中国法律评论》2018年第6期,第90-102页;唐应茂:《司法公开及其决定因素:基于中国裁判文书网的数据分析》,载《清华法学》2018年第4期,第35-47页。

[37] 杨金晶、覃慧、何海波:《裁判文书上网公开的中国实践——进展、问题与完善》,载《中国法律评论》2019年第6期,第133-135页。

[38] 《最高人民法院关于人民法院在互联网公布裁判文书的规定》(法释〔2016〕19号),第16条。

开⁽³⁹⁾；敏感事件可能也会导致裁判文书网下线部分文书。⁽⁴⁰⁾裁判文书的总撤回比例仍是未知数。

再次，裁判文书中的内容不是司法实践的完全体现。裁判文书内容往往是对事实、主张和裁判结果的简要概括总结。司法决策过程中的博弈过程、裁判目的、心证形成、内部讨论等影响决策的关键因素往往是高度非文字化、非数据化的，无法记录甚至不允许记录，不能体现在裁判文书中。⁽⁴¹⁾裁判文书内容表现出的论证说理与法官实际的裁判思维之间也可能存在差异。⁽⁴²⁾裁判文书内容的质量存在显著差异，囿于不同级别、不同地域裁判文书的行文风格以及文书质量等客观因素，裁判文书之间体现出的差异可能是由于裁判文书质量不同，而非真实裁判倾向的不同，即裁判文书数据可能无法直接准确地说明全国法院的裁判倾向。

最后是信息抽取的局限性。面对裁判文书中的多样化表述，利用机器学习技术进行信息抽取仍然面临挑战，存在改进空间。⁽⁴³⁾

简而言之，不是所有的法律纠纷都会通过诉讼解决，也不是所有的司法判决的裁判文书都会公开；裁判文书没有记载司法实践中的全部信息，裁判文书中记载的信息目前也无法完全准确地抽取出来。因此，研究者可获取的裁判文书和信息相对全体样本可能存在一定程度的错位，即存在选择性偏差⁽⁴⁴⁾等局限性。虽然裁判文书数据体现了司法实践中的某些现象，但不能不加分析地直接断言裁判文书数据中体现的现象对所有知假买假行为具有足够的代表性。

（五）对裁判文书数据局限性的应对思路

裁判文书数据的局限性不能完全消除，但这并不意味着这些数据一无是处，裁判文书数据仍然是法实证研究的重要资源。法实证研究不必奢求数据的完全准确，重点在于利用归纳论证的盖然性思维，通过实证材料（数据）为法学研究中的事实判断问题给出有更高盖然性的结论。

归纳论证并不能对结论作确定性的断言，其结论是否为真只是具有某个概率度。评价归纳论证的强弱标准就在于其结论的盖然性程度的高低。⁽⁴⁵⁾对事实判断问题的回答结论正如现代科学理

〔39〕 徐文进、姚竞燕：《被遗忘权范式下裁判文书上网后撤回机制的检视与优化——基于 131 份撤回文书及〈个人信息保护法〉的定向分析》，载《法治研究》2022 年第 1 期，第 82—90 页。

〔40〕 例如，笔者于 2022 年 2 月 19 日在裁判文书网以“收买被拐卖的妇女罪”为关键词，检索案件类型为“刑事案件”，文书类型为“判决书”的文书，结果为 477 篇；但于 2022 年 4 月 4 日以同样的条件检索，结果仅为 421 篇。

〔41〕 左卫民：《关于法律人工智能在中国运用前景的若干思考》，载《清华法学》2018 年第 2 期，第 115 页。

〔42〕 由于法官精力所限或避免担责等因素，部分法官不愿过多说理，而是仅仅简单、抽象地概括。参见庄绪龙：《裁判文书“说理难”的现实语境与制度理性》，载《法律适用》2015 年第 11 期，第 83—87 页。法官在裁判文书中的说理可能仅仅是以既有的法律规定作为表面的论证理由，掩饰自己内心真实的裁判理由和推理过程。参见李学尧、刘庄：《矫饰的技术：司法说理与判决中的偏见》，载《中国法律评论》2022 年第 2 期，第 91—106 页。

〔43〕 在最近的一项法律事件探测研究中，研究者基于前所未有的大规模标注数据、利用优化后的算法得到的法律事件多分类模型的 micro-F1 值为 85.48%，仍须进一步提高。Feng Yao et al., *LEVEN: A Large-Scale Chinese Legal Event Detection Dataset*, Findings of the Association for Computational Linguistics: ACL 2022, 188 (2022).

〔44〕 选择性偏差 (selection bias)，也称为选择性效应 (selection effect)，是指在研究过程中因样本选择的非随机性而导致得到的结论不能代表总体特征。产生选择性偏差的典型情景如“去公园调查居民健康水平”。

〔45〕 [美] 欧文·M·柯匹、[美] 卡尔·科恩：《逻辑学导论》(第 13 版)，张建军等译，中国人民大学出版社 2014 年版，第 627 页。

论及其包括的因果律一样，都仅仅是盖然的。所有探索世界的现代科学命题本质上都仅仅只是尚未被证伪的假说，不具有绝对的确定性，并不断在经验中检验和发展。^[46]如果把法实证研究视为科学研究，那么法实证研究不提供确定性的结论，而只是尽可能提高结论的盖然性。

法的实际样态是一种复杂的社会现象，对它的每一种观察都难以做到周全，一个研究者不可能穷尽所有相关材料。法实证研究不提供确定性的教条，而是在正视自身的局限性的同时，以开放的态度欢迎对旧成果的检验、欢迎对新的材料和方法的使用。即使检验或基于新的材料、方法得出的结论与旧结论相同，也只是加固了旧结论，而非为旧结论提供了确定性；如果检验或基于新的材料、方法得出的结论与旧结论存在差异，也可能只是削弱了旧结论。但在通过检验或新的材料、方法将旧结论证伪之前，基于不完全数据的研究仍具有价值：面对法学中的事实判断问题，至少系统的实证研究提供了比个别经验或主观推断有更高的盖然性的回答。

二、研究的开始：法学的问题意识及其拆解

实证研究从问题开始。“所谓的实证研究，不是以方法来界定，或者是以概念来界定，而是以真正的问题来界定。要看这个问题是不是有意思，是不是我们必须面对的问题。”^[47]有时甚至可以说，“重要的是问一个与法律有关、能够量化进行研究的问题”。^[48]

（一）问题意识的确立

问题意识包括两个方面：法学规范问题与可检验的事实判断问题。^[49]可检验的事实判断问题以法学规范问题为引导，是研究法学规范问题的中介。通过澄清事实前提，对相互冲突的竞争性理论解释方案进行结果论或目的论的验证和筛选是法实证研究问题意识的重要来源之一。

没有问题意识指导的数据分析意义不大。问题意识决定了选用何种实证材料、抽取实证材料中的何种信息以及分析数据的方式。^[50]而在接触法律现象、形成问题意识并确定研究的对象、内容和方法时，人们又总是带着一定的“前理解”或预设。这些“前理解”或预设包含着研究者对法学问题和方法问题的思考，并且总是在对关于法与社会的一般理论的整体把握与对研究对象的具体认识之间的关联中形成的。因此，可以说法实证研究是在法学的规范问题与方法问题、观察与理解交织形成的“分析框架”下开展的。^[51]

[46] [美] 欧文·M·柯匹、[美] 卡尔·科恩：《逻辑学导论》（第13版），张建军等译，中国人民大学出版社2014年版，第595-598页。

[47] 苏力：《好的研究与实证研究》，载《法学》2013年第4期，第17页。

[48] 张永健、程金华：《法律实证研究的方法坐标》，载《中国法律评论》2018年第6期，第77页。

[49] 这种可检验的事实判断问题可以是某种假设或理论，表现为任何关于什么导致什么、两个变量之间为何且如何相关的观点。参见[美] Donald J. Treiman（唐启明）：《量化数据分析：通过社会研究检验想法》，任强译，社会科学文献出版社2012年版，第26页。

[50] 关于事实判断问题如何决定实证材料的选取和处理方法，本文没有专门讨论，这要视研究目的、研究者的资源等具体情形而定。例如，如果要讨论司法实践中的有关事实，那么裁判文书是重要的实证材料；如果做事实判断需要每一份裁判文书中具体的信息要素，那就需要对裁判文书进行深度的信息抽取。

[51] 关于“分析框架”的讨论，参见彭小龙：《法实证研究中的“理论”问题》，载《法制与社会发展》2022年第4期，第211页。

以“知假买假”问题为例，消费者惩罚性赔偿制度⁽⁵²⁾的规范目的是鼓励消费者对不法商家提起诉讼，填补监管空白，从而惩罚和遏制不法行为，保护消费者利益。⁽⁵³⁾但我国的消费者惩罚性赔偿制度诞生后，出现了利用该制度牟利的知假买假行为。在个案中进行价值判断虽然有助于实现个案正义，但无法直接导向制度性改进。⁽⁵⁴⁾围绕“知假买假者是否应当获得惩罚性赔偿金”这一制度问题，二十余年来学术界和实务界在道德判断、法律概念体系等论证维度争论不休，但没有得出被广泛接受的结论。⁽⁵⁵⁾关于知假买假问题，最高人民法院曾先后多次分别发布过司法解释、指导性案例和工作答复，部分地方法院也发布过相关意见。⁽⁵⁶⁾但是，由于学术界、司法界等各界对知假买假相关的事实判断问题没有明确的认识，这些相互之间存在差异的文件并没有终结关于知假买假问题的争议。

道德判断和法律概念体系的论争似乎进入了死胡同，近来关于知假买假的研究逐渐在论证维度上转向注重功能主义。⁽⁵⁷⁾法律规范有效力并不意味着其预设目标当然地得到实现。功能主义怀疑纯粹的道德判断或概念体系分析在知假买假这类疑难法律问题中的论证有效性，转而强调就法律的不同形式化解释方案对行为选择和社会秩序产生的现实影响进行观测和比较，从而选择那些更有助于实现特定政策目标的法律适用方案。⁽⁵⁸⁾最高人民法院近来对知假买假问题的处理也体现

(52) “消费者惩罚性赔偿制度”，是指购买者根据《消费者权益保护法》（1993年通过）第49条、《消费者权益保护法》（2013年修正）第55条第1款、《食品安全法》（2009年通过）第96条第2款、《食品安全法》（2015年修订）第148条第2款或《药品管理法》（2019年修订）第144条第3款请求商品生产者、经营者，服务提供者或药品生产企业、经营企业、医疗机构支付惩罚性赔偿金的制度。

(53) 王利明：《惩罚性赔偿研究》，载《中国社会科学》2000年第4期，第115-116页。

(54) 在既有制度条件下，面对诸如知假买假能否请求惩罚性赔偿等疑难问题的考问，概念法学已然无法应对自如，此时需要运用价值补充方法进行裁断，权衡具体的打假行为是否有助于发挥制度本身的作用与价值。参见姚辉：《民法适用中的价值判断》，载《中国法律评论》2019年第3期，第93-112页。

(55) 道德判断的理由诸如“不劳而获”“以暴制暴”等，法律概念体系的理由诸如“知假买假者不属于‘消费者’”等。单纯的道德判断或形式逻辑推演容易陷入无休止的价值判断或形式化对立困境。对相关实务及学术争议的评析参见熊丙万：《法律的形式与功能——以“知假买假”案为分析范例》，载《中外法学》2017年第2期，第300-339页。

(56) 《最高人民法院关于审理食品药品纠纷案件适用法律若干问题的规定》（法释〔2013〕28号）第3条规定：“因食品、药品质量问题发生纠纷，购买者向生产者、销售者主张权利，生产者、销售者以购买者明知食品、药品存在质量问题而仍然购买为由进行抗辩的，人民法院不予支持。”（该司法解释经过2020年和2021年两次修订，现行有效版本的文号为法释〔2021〕17号，但其中第3条自2013年以来未发生变动。）指导性案例是指孙某某诉南京欧尚超市有限公司江宁店买卖合同纠纷案，最高人民法院指导性案例23号（2014年），支持知假买假者的惩罚性赔偿请求权。工作答复是指《最高人民法院办公厅对十二届全国人大五次会议第5990号建议的答复意见》（法办函〔2017〕181号），明确反对非食药领域的知假买假者的惩罚性赔偿请求权。地方法院的意见例如2014年5月《深圳市中级人民法院关于消费者权益纠纷案件的裁判指引》第1条明确认可知假买假者属于消费者；2016年3月《重庆市高级人民法院关于审理消费者权益保护纠纷案件若干问题的解答》（渝高法〔2016〕77号）第2条明确反对知假买假者的惩罚性赔偿请求权。

(57) 葛江虹：《“知假买假”：基于功能主义的评价标准构建与实践应用》，载《法学家》2020年第1期，第160-172、196页。关于在民法学研究中坚持功能主义思维的一般性分析，参见申晨：《论中国民法学研究中的功能主义范式》，载《法制与社会发展》2023年第5期，第146-166页；许中缘：《论民法典功能主义释意与民法教义学的重构》，载《当代法学》2023年第5期，第19-31页。

(58) 虽然法律需要以一套概念术语及其体系结构获得外在表达，以便于被人们认识、交流和遵守，但这些形式背后却有着明显的功能预设，承载了实现特定的理想社会秩序之功能。而法律的概念体系是对复杂现实生活进行过滤和抽象的产物。不同的形式化解释和适用方案又会对社会现实生活产生不同的功能。参见熊丙万：《法律的形式与功能——以“知假买假”案为分析范例》，载《中外法学》2017年第2期，第314页。

出了一定的功能主义倾向。2023年11月30日发布的《最高人民法院关于审理食品药品惩罚性赔偿纠纷案件适用法律若干问题的解释（征求意见稿）》在明确不能以“知道所购买食品不符合食品安全标准仍然购买”否定惩罚性赔偿金请求的同时，拟通过“合理生活消费需要范围”限制计算惩罚性赔偿金的基数，并以食品自身的安全标准作为认定惩罚性赔偿责任的依据。这些规定在一定程度上区分了不同性质的知假买假行为，但能否实现相应的政策目标还有待进一步评估。

功能主义强调的是，先回答“允许知假买假者获得惩罚性赔偿金对社会秩序产生的现实影响”这一事实判断问题，再将其结论用于规范论证。由此，便提出了一个需要进行实证研究的法学问题，即由“知假买假者是否应当获得惩罚性赔偿金”问题衍生出的“允许知假买假者获得惩罚性赔偿金的现实影响为何”。具体而言，从目的论证的角度，需要回答的事实判断问题是“如果说消费者惩罚性赔偿制度预设的目的是惩罚和遏制不法行为、净化市场环境、保护消费者利益，那么，允许知假买假者获得惩罚性赔偿金是否有利于这个目的的实现？”（下称目的实现问题）。从结果论证的角度，需要回答的事实判断问题是“允许知假买假者获得惩罚性赔偿金会导致何种结果？”（下称结果判断问题）。

探究上述事实判断问题需要获得关于知假买假行为的信息，而裁判文书中会记录法院查明的事实，其中会包含知假买假行为事实，因而裁判文书成了知假买假法实证研究的重要材料。

（二）问题与概念的拆解

截至本文发表时，已公开的裁判文书中涉及消费者惩罚性赔偿条款的约有10万份，但裁判文书文本并不能直接开口告诉我们事实判断问题的答案。因此，需要将比较宏观的事实判断问题拆解为较为具体的问题，并最终对应到能够抽取出的裁判文书中的“要素”。^{〔59〕}问题与概念的拆解就是将抽象的法律概念与裁判文书信息对应的操作过程，也是联系理论世界与经验世界的过程。

篇幅所限，此处仅以目的实现问题中的一个方面举例。若要分析允许知假买假者获得惩罚性赔偿金是否有利于净化市场环境，一般统计学上的因果推断需要不同制度条件下的市场环境指标等数据。但是，制度条件和市场环境难以量化，市场环境的影响因素也过于复杂。裁判文书能体现的事实信息主要是原被告双方的行为，不妨直接分析知假买假者的行为多大程度上属于有利于净化市场环境的行为。具体而言，可以至少从以下几个方面考虑：（1）知假买假者的起诉理由多大程度上是实质性的安全或误导问题？（2）知假买假者起诉针对的商品多大程度上被法院认定为确实存在问题？（3）知假买假者是否以及多大程度上存在反复多次起诉的行为？（4）知假买假者起诉的商家中正规性较高的大商超、大厂家和监管相对薄弱的小卖店、小厂家各自占比如何？^{〔60〕}

〔59〕 这里的“要素”是指可以从文书中直接提取的最小颗粒度的信息项，包括裁判文书中的事实、主张、论证、裁判等各方面的最小颗粒度信息。这与要素式审判中的所指要素有所不同，要素式审判中的要素一般指能够影响法官裁量的最小颗粒度的有法律意义的事实信息。关于要素式审判的讨论参见林遥：《民事类型化案件要素式审判机制研究——以C市法院民事庭审优质化改革情况为样本分析》，载《法律适用》2018年第15期，第135-142页。

〔60〕 例如，可根据被告商家主体性质进行区分。被告为自然人或个体工商户的为一类，被告为公司的再以其注册资本（被告为公司或分店的，以其所属总公司的注册资本计算）为标准区分“大”和“小”，分析被告注册资本的分布。

(5) 被起诉的商家在败诉后是否会因同类事由再次被起诉(即依然存在问题)?⁽⁶¹⁾为回答这些具体问题,需要确定通过裁判文书能体现什么信息。判断的方式不一定是单独一份文书,也可以是不同文书间的比较。

针对问题一,需要考察知假买假者作为原告起诉时的事实和理由,并依原告的主张区分为不同的打假事项。实质性的安全问题诸如过期食品、含有有毒有害物质、无照生产等;实质性的误导问题诸如夸大功效、夸大品质、虚构评价、虚构荣誉等;而非实质性的问题诸如标签说明书存在错别字等不影响安全且不会造成误导的瑕疵、仅违反行政管理规定等。同时,还可以根据出现的不同问题,对打假作类型区分,包括有毒有害、误导决策、轻微行政违规、碰瓷四类。

针对问题二,需要考察法院在事实认定和裁判说理中对涉案商品问题的认定情况进行分类,可以分为存在/不存在安全问题、会/不会误导消费决策、违反/不违反行政管理规定。当然,上述几种情况也可以简略地分为“有问题”和“没问题”,但细致的分类并非没有法律或学术意义⁽⁶²⁾,为了研究的可拓展性,在信息抽取时进行相对细致的分类是更恰当的做法。

针对问题三,需要每一份裁判文书中的原告信息、被告信息,以及涉案商品信息等案件事实信息,综合判断是否构成重复起诉。

针对问题四,需要每一份裁判文书中的被告信息,必要时需要利用被告主体信息在外部数据库中查询企业注册资本等工商信息,判断被告的性质。

针对问题五,需要比较不同裁判文书之间的被告信息、涉案商品信息、打假事项信息、判决结果信息。

除了针对具体问题的信息外,为了在时间、空间维度进行分析,还需要法院所在地、审判程序、审判时间等一般性的案件基本信息。

问题与概念的拆解不是纯粹事实判断,上述拆解和分类方式多有理论背景;这种拆解也不是一蹴而就的,拆解的过程需要“在实证资料、学说资源、问题意识之间目光往返流转”⁽⁶³⁾,动态地深化对问题、材料、方法的理解。在研究开始时,问题意识并不一定特别具体、清晰,对于材料的理解也会有所欠缺,因而问题和概念的拆解也会存在不完善之处,比如某个问题被遗漏、某项信息被遗漏或重复、某些信息在裁判文书中无法体现,抑或是拆解本身存在问题;在研究过程中也可能发现某些问题无法验证。

例如,基于功能主义分析,有必要区分哪些知假买假行为会产生积极或消极的现实影响,进而将打假类型区分为良性打假与恶性打假两类。⁽⁶⁴⁾但是,这种二分法难以充分涵盖复杂的知假买

(61) 当然,这些问题的成立包含着前提假设,如第一个问题包含着如下前提:“知假买假者的起诉理由越是针对实质性的安全或误导问题,就越有利于净化市场环境。”有时也需要检验这些前提成立与否。

(62) 比如,因安全问题和欺诈提起诉讼的请求权基础不同,安全问题和欺诈的认定标准也不同。

(63) 彭小龙:《法实证研究中的“理论”问题》,载《法制与社会发展》2022年第4期,第218-220页。

(64) 熊丙万:《法律的形式与功能——以“知假买假”案为分析范例》,载《中外法学》2017年第2期,第338-339页。

假现象，也对惩罚性赔偿制度的功能发挥作了过于简化的理解。^{〔65〕}通过对相关裁判文书进一步阅读也可以发现，法官对知假买假行为的判断并不是良性或恶性的二分法，而是更加关注其所针对的商品存在的问题究竟是安全问题、欺诈问题还是轻微行政违规等问题。因此，可以根据打假行为所针对的商品问题类别将打假类型拆解为有毒有害、误导决策、轻微行政违规和碰瓷四类。

又如，有学者认为知假买假者起诉时不能完全确定法院是否会支持其惩罚性赔偿请求，带有投机和侥幸心理。^{〔66〕}起诉时不能确定裁判结果是否意味着投机和侥幸心理，进而作否定评价，自有待商榷。但这一学说资源提示了一个问题，即“知假买假行为在多大程度上具有投机性”。裁断是否“投机”属于主观动机，也涉及价值评价，很难在裁判文书这样的实证材料中体现出来。裁判文书数据可以回答的事实判断问题是前述问题二，即“知假买假者起诉针对的商品多大程度上被法院认定为确实存在问题？”

三、研究的关键：裁判文书信息抽取的技术操作

经过问题意识的确立和拆解，就大致得到了有待回答的事实判断问题和裁判文书中具体要素的对应关系。但是，裁判文书信息抽取需要专门的技术处理。本节以大规模消费者惩罚性赔偿裁判文书的信息抽取过程为例，系统探讨利用机器学习等技术进行裁判文书信息抽取的实现步骤。裁判文书信息抽取的核心是在正确进行人工标注的前提下，使得计算机算法能够最大程度上与人工标注方式拟合，从而让计算机能够像人一样标注裁判文书，由此，便可获知每一份裁判文书中的要素信息。为实现这一目标，裁判文书信息抽取的流程如图2所示，后文详细解释。

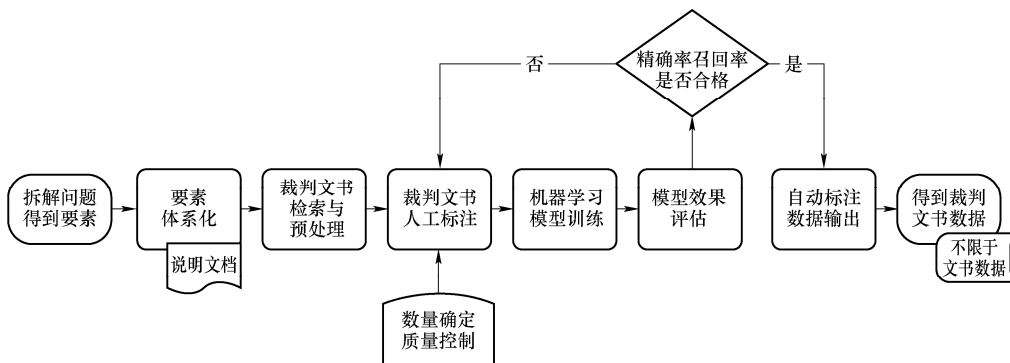


图2 裁判文书信息抽取流程

（一）要素的体系化

现有的法实证研究抽取的裁判文书信息往往层次较浅、数量较少。信息层次多为案由分布、法院层级、诉讼主体身份、有无律师、文书引用的法条等浅层信息，而对于原被告的主张、法院

〔65〕 葛江虬：《“知假买假”：基于功能主义的评价标准构建与实践应用》，载《法学家》2020年第1期，第169-172页。

〔66〕 该文认为，知假买假行为与赌博有相似之处。比照刑法中的赌博罪，知假买假不值得被鼓励，“形成组织或者以买假索赔为业的，不属于消费者”。参见刘保玉、魏振华：《“知假买假”的理论阐释与法律适用》，载《法学论坛》2017年第3期，第66页。

的说理和裁判结果等主文的分析则相对较少。⁽⁶⁷⁾挖掘的信息层次浅也就决定了这些研究中抽取的要素数量较少。而在知假买假问题的法实证研究中,研究开始时问题与概念的拆解过程中已经预设了需要从裁判文书中抽取出种类繁多的要素信息,一个问题需要诸多不同层次、不同类别的要素信息,不同的问题也可能需要相同或有包含关系的要素。因此,为了提高裁判文书信息抽取的效率,需要根据裁判文书自身的写作逻辑和结构将拆解出的要素组织起来,如将涉及原告主张的、事实认定的、法官论证的要素分别归类。体系化的目标是形成一个适应裁判文书结构的层次分明、逻辑清晰、结构明确的概念与要素体系,避免要素之间的重复或矛盾。在要素体系化的过程中,需要编写与要素体系相同结构的要素说明文档,对每一个要素的含义与标注标准作出说明。编写要素说明是保障工作的延续性与知识共享性的必要工作,也是人工标注与结果校验的指南。

通过前述问题与概念的拆解,共形成了192个要素,经过要素的体系化,这些要素组成了一个多层次的树状结构。例如,“维护消费者权益”“维护商标权人利益”“促进合法经营”“多元治理”四个要素属于对知假买假基于功能主义的支持理由;功能主义的支持理由又与功能主义的反对理由、道德主义的支持与反对理由、形式主义各自的支持与反对理由共同归类为法官的论证方向;法官的论证方向与法官的主客观审查、裁判结果构成文书的裁判情况;裁判情况与案件基本信息、原告信息、被告信息、商品信息、案件事实共同构成要素的六大分类。

(二) 裁判文书的检索与预处理

裁判文书的检索旨在从裁判文书全库中准确挑选出全部符合研究需要的裁判文书。目前,比较通行的做法是与法律科技公司、法律数据公司合作,在它们的裁判文书全库中检索并输出研究需要的裁判文书。消费者惩罚性赔偿这类诉讼的请求权基础是确定的⁽⁶⁸⁾,且这些法条基本只对应消费者惩罚性赔偿诉讼,故可以以文书引用的法条作为检索条件。2020年初检索得到的惩罚性赔偿裁判文书数为6万余份,截至本文发表时已经增加到约10万份,并还在持续增加中。如果是检索某些行为方式多样、案由多样的案件的裁判文书,可能就需要设计比较复杂的检索方案。

检索得到的裁判文书中存在非标准化文书,在进行下一步工作之前要先对裁判文书进行预处理。预处理的目的是提高文本数据质量、方便后续工作。裁判文书的预处理主要包括如下方面。

(1) 文书去重。有时在裁判文书网上内容相同的裁判文书会重复出现。对于重复文书,相同案号的文书可以只保留一份;或根据文书内容去重,内容完全相同的文书只保留一份。

(2) 数据清洗。⁽⁶⁹⁾虽然裁判文书是格式化程度较高的法律文本,但仍有相当数量的文书存在各种各样的写作错误。⁽⁷⁰⁾识别这些错误的方法主要有:①对于每一份裁判文书都应当有的要素

(67) 左卫民、王婵媛:《基于裁判文书网的大数据法律研究:反思与前瞻》,载《华东政法大学学报》2020年第2期,第69页。

(68) 《消费者权益保护法》(1993年通过)第49条、《消费者权益保护法》(2013年修正)第55条第1款、《食品安全法》(2009年通过)第96条第2款、《食品安全法》(2015年修订)第148条第2款或《药品管理法》(2019年修订)第144条第3款。

(69) 数据清洗是通过修改数据的形式或内容来提高数据质量的过程。Ron Kohavi & Foster Provost, *Glossary of Terms*, 30 *Machine Learning* 271, 272 (1998)。

(70) 马超、于晓虹、何海波:《大数据分析:中国司法裁判文书上网公开报告》,载《中国法律评论》2016年第4期,第199、240-241页。

(如原告诉请、审判结果等)未输出数据;②同一份裁判文书输出的数据相互矛盾(如既认定商品不存在问题,又支持惩罚性赔偿金的诉请);③人工对常见写作错误的经验总结。现阶段还不能通过自动化方式识别出全部错误。识别出错误后,对于能够挽回的数据,可以进行人工修正。

(3)裁判文书内容分段。^[71]分段的原因在于,某个表述(可能关于事实、证据或法律适用)可能会被原告、被告、法官多次提及,同类表述在不同的段落出现可能意味着不同的信息。分段可以使信息抽取的结果更准确。

(三)裁判文书信息抽取的方式选择

一般的关键词匹配通常无法满足裁判文书信息抽取的需要。实践中,裁判文书信息抽取主要有三种方式:正则表达式、人工标注、机器学习。其中,正则表达式^[72]以对要素对应的裁判文书语句的人工总结为基础,需要人工反复测试和修改完善。裁判文书人工标注以前述要素体系的构建为前提,通过人工标注可以得到每一要素与裁判文书中语句的对应关系。标注结果既可以直接导出为结构化数据,也可以为机器学习^[73]提供训练样本。人工标注能处理的文书数量是有限的,而正则表达式和机器学习模型都可以自动复用到新增文书上。

选择何种方式取决于要素信息在裁判文书中的特征。在惩罚性赔偿裁判文书的信息抽取工作中,根据不同的要素信息在裁判文书中表述特征的不同,灵活选用了不同的抽取方式。案件基本信息的信息抽取主要使用正则表达式^[74],原告信息、被告信息、案件事实、法官的论证等方面的信息抽取主要使用人工标注+机器学习。

(四)对裁判文书的人工标注

人工智能从来是有多少人工就有多少智能。尤其是在监督学习中,如果没有人类为数据打标签,机器将无法进行学习。虽然有研究者认为人工打标签效率低、限制了学习样本量^[75],但目前缺乏更好的解决方案,人工标注仍然是必要的。

1. 标注数量的确定

人工标注以文书为单位进行,机器学习的模型训练(后文详述)以有标注样本(语句)为单位进行。针对每个要素的机器学习任务都可以理解为判断一句话是否表述了该要素信息的文本二

[71] 例如,一份标准的民事判决书由标题、正文、落款三部分组成,其中正文包括首部、事实、理由、裁判依据、裁判主文、尾部;正文中的事实又包括原告的诉讼请求、事实和理由,被告答辩的事实和理由,法院认定的事实和据以定案的证据。关于裁判文书的分段标准,参见《最高人民法院关于印发〈人民法院民事裁判文书制作规范〉〈民事诉讼文书样式〉的通知》(法〔2016〕221号)。不同部分的分界处一般存在比较确定的标志性词汇或短语;但需要注意,案情或法官固有写作习惯等的差异会导致文书之间存在一些差别。

[72] 正则表达式可以大致理解为能够在关键词的基础上设置复杂条件限定与条件嵌套的文本匹配方法。关于使用正则表达式进行裁判文书信息抽取,参见邓矜婷、张建悦:《计算法学:作为一种新的法学研究方法》,载《法学》2019年第4期,第115-121页;周翔:《作为法学研究方法的大数据技术》,载《法学家》2021年第6期,第66-67页。

[73] 机器学习能够通过算法使计算机从历史数据中学习规律,并利用经验来改善系统自身的性能,从而对新的样本进行识别。参见周志华:《机器学习》,清华大学出版社2016年版,第1页。

[74] 如案号、案由、法院所在地、裁判日期等要素使用正则表达式进行抽取的准确率在95%以上。

[75] 王禄生:《司法大数据与人工智能开发的技术障碍》,载《中国法律评论》2018年第2期,第51页。

元分类任务。对机器学习来说,一般训练样本越多就越有可能得到效果好的模型。⁽⁷⁶⁾ 一项法律事件检测 (Legal Event Detection, LED) 研究显示:总体来说,模型性能随着训练实例数量的减少而下降;实践显示,至少需要 150 个训练实例才能使得模型不因实例数量过少而得出较差结果。⁽⁷⁷⁾ 但是人力有限,标注样本量不可能无限增多。综合考虑各要素在文书中的出现频率以及标注成本,惩罚性赔偿裁判文书信息抽取工作确定了约 6 000 份文书的标注数量。⁽⁷⁸⁾ 为了保证数据代表性,该约 6 000 份文书应当在检索到的全部相关文书中随机抽取。

经过人工标注,并删除其中的不相关文书以及标注质量较差的文书,最终得到了 5 406 份有标注惩罚性赔偿的裁判文书,101 262 条标注数据。

2. 标注质量的控制

机器学习的模型效果高度依赖于用以训练模型的输入数据的质量,遵循“垃圾进,垃圾出”(garbage in, garbage out) 的定律。因此,标注质量对于模型效果及基于模型的自动标注的数据质量至关重要。

人工智能是没有理解的“智能”。机器学习能学到的并非语句的含义,而是符号的组合方式和对应关系。因此,人工标注的语句应当是表意明确,不需要推理即可识别意思的语句。⁽⁷⁹⁾ 文本标注中的保障标注数据质量的方法主要包括:(1) 基于要素说明文档对标注人员进行培训,并通过试标注检验标注人员对要素内涵理解的准确性及一致性;(2) 每一标注人员在一段时间内仅针对一类要素进行标注,以节约注意力资源避免遗漏并提高标注效率;(3) 每一份文书由不同的标注人员独立进行两次标注,并进行标注一致性检验⁽⁸⁰⁾,若两人的标注一致性较低,则将文书随机分配给第三人进行标注,若无法形成多数意见则记为疑难文书单独处理;(4) 为避免标注者之间共性的理解误差,标注完成后对结果随机核查⁽⁸¹⁾; (5) 视情况对部分标注效果较差的要素进行删除

(76) 周志华:《机器学习》,清华大学出版社 2016 年版,第 3-4 页。

(77) Feng Yao et al., *LEVEN: A Large-Scale Chinese Legal Event Detection Dataset*, Findings of the Association for Computational Linguistics: ACL 2022, 189 (2022).

(78) 150 个实例相对 6 000 份文书仅占 2.5%,对 6 000 份文书进行人工标注可以使大多数要素实例数量满足机器学习文本分类任务的需要。

(79) 例如在广东省深圳市龙岗区人民法院民事判决书(2014)深龙法坂民初字第 376 号中,读者可以从判决理由段的“被告作为食品销售者,负有保证食品安全的法定义务,应当对不符合安全标准的食品及时清理下架”一句判断出被告销售了不符合食品安全标准的食品。但是,这种判断存在推断成分。而这份文书中的另一句“但被告仍然进行销售,系不履行法定义务的行为,应当被认定为销售明知是不符合食品安全标准的食品”,则不需要推断即可看出被告销售了不符合食品安全标准的食品。因此,在这份文书中应当标注为“存在安全问题”要素的语句是后者而不是前者。

(80) 检验标注一致性的指标一般是 Kappa 系数或 Alpha 系数,这些系数是一个系列,根据任务种类的不同有不同的具体计算方式。最近的一项裁判文书中法律事件检测研究中,研究者对裁判文书进行了双重标注并使用了 Cohen's Kappa 系数评估标注了一致性。参见 Feng Yao et al., *LEVEN: A Large-Scale Chinese Legal Event Detection Dataset*, Findings of the Association for Computational Linguistics: ACL 2022, 187 (2022).

(81) 当要素的标注比例未知、裁判文书总量大于 385 时,抽取的样本容量和总体的数量无关,以有放回抽样法进行人工抽样检查的样本容量可以恒为 385,此时有 95% 的把握认为样本的准确率和总量准确率的误差不超过 5%。参见贾俊平编著:《统计学基础》(第 2 版),中国人民大学出版社 2014 年版,第 137 页。

处理或对部分要素重新标注。

惩罚性赔偿裁判文书的信息抽取基本遵循了上述标注质量控制方法。但标注一致性检验并没有采用双重标注方法，而是利用了“拆分诉讼”⁽⁸²⁾ 文书的标注结果。拆分诉讼文书的共同特征是同一组案件的文书内容高度相似。每一组案件，少的拆分为两起诉讼，多的则拆分为数十起乃至上百起诉讼。对于每一组拆分诉讼案件，可以视为对一起案件的多重标注。因此，可以通过对拆分诉讼案件的标注一致性检验推断整体的标注一致性。Fleiss's Kappa 系数可以用来衡量多标注者多要素的标注一致性。⁽⁸³⁾ 经初步核验，这些文书的整体标注一致性较好，但个别要素需要改进。

(五) 基于人工标注的机器学习模型训练

虽然法官在个案中的价值判断存在主观因素，但法官的价值判断在裁判文书中的表述存在客观规律。⁽⁸⁴⁾ 具体来说，裁判文书信息抽取中机器学习的任务是通过算法寻找被人工标注了某个要素标签的所有语句在表述上的某种规律，以构建裁判文书语句和要素标签之间的对应关系（模型训练）。此处的“模型”描述的仅仅是裁判文书中的语句和某个要素之间的对应关系，计算机获得的仅仅是自动判断某一语句是否是在表述某个特定要素的能力。每一个要素在技术上都是判断裁判文书中某个语句是否表达了该要素信息的二分类任务。

机器学习的模型训练工作主要是计算机科学领域的任务。法学研究者不必过于关心模型训练的具体技术细节，如使用了什么算法、如何改进算法等。实际上，目前的几种前沿算法在处理裁判文书文本时的效果相近，不存在根本性区别。⁽⁸⁵⁾

(六) 模型效果评估

对于自动标注的二分类任务，不必过度关注模型的可解释性，而是应主要关注模型输出结果是否以及在多大程度上存在错漏。评估模型效果的基本思路是将标注出的样本数据分为两部分：训练集和测试集。其中，训练集用来训练模型，测试集用来测试模型对新样本的判别能力，测试集应尽可能与训练集互斥。划分训练集和测试集的方式主要有“留出法”“交叉验证法”“自助法”等，其中“交叉验证法”“自助法”会产生多组不同的数据集划分结果。⁽⁸⁶⁾

不论如何划分，都可将训练集上的模型预测结果和人工标注结果进行对比，得到如表 1 所示的二分类任务混淆矩阵。

〔82〕 拆分诉讼案件，即一名商品购买者在同一经营者处购买多件商品、分别结账并分别提起惩罚性赔偿诉讼。对于这类案件，法官一般会共用一套论证方式，裁判文书内容基本一致。之所以没有在一开始想到利用拆分诉讼文书，是因为拆分诉讼现象是研究过程中逐渐发现的。这体现了实证研究对于发现新的事实的价值。

〔83〕 Artstein R, *Inter-Annotator Agreement*, in Nancy Ide & James Pustejovsky eds., *Handbook of Linguistic Annotation*, Dordrecht Springer, 2017, p. 297–313.

〔84〕 季卫东：《计算法学的疆域》，载《社会科学辑刊》2021年第3期，第121页。

〔85〕 Feng Yao et al., *LEVEN: A Large-Scale Chinese Legal Event Detection Dataset*, *Findings of the Association for Computational Linguistics: ACL 2022*, 188 (2022).

〔86〕 周志华：《机器学习》，清华大学出版社2016年版，第23–27页。

表 1 二分类任务混淆矩阵

		模型预测值	
		1 (Positive)	0 (Negative)
人工标注值	1 (Positive)	TP (True Positive)	FN (False Negative)
	0 (Negative)	FP (False Positive)	TN (True Negative)

评估一个模型效果的主要有精确率 (Precision, P)、召回率 (Recall, R)、 F_1 值 (F_1 -score, F_1) 三个指标。其中, 精确率衡量结果的正确性, 召回率衡量结果的全面性, F_1 值综合衡量正确性和全面性。基于表 1, 上述三个模型效果指标的计算方式如下:

$$P = TP / (TP + FP)$$

$$R = TP / (TP + FN)$$

$$F_1 = (2 \times P \times R) / (P + R)$$

对裁判文书的自动标注应当既正确又全面, 即有较高的 F_1 值。对于定量分析而言, 如果模型 F_1 值能够达到 0.95 甚至以上的标准, 的确是十分理想的结果。⁽⁸⁷⁾ 但对于不少要素的标注而言, 这一点实现起来很有挑战性, 有必要予以适度调整。对于那些低频要素, 即使文书数量很多, 实例数量也过少, 影响了机器学习效果, 可能难以满足基本的定量分析要求。同时, 对于那些带有主观价值判断特点的要素而言, 即便实例数量足够丰富, 模型 F_1 值也可能因为主观表达方式的多样性而相对较低。例如在惩罚性赔偿裁判文书的信息抽取中, 要素“法官的论证/道德主义-反对/有违诚信原则”的模型 F_1 值可达 90.1%。这表明, 一方面, 通过机器学习获得的可用数据的质量, 还有通过进一步努力来提升的空间。另一方面, F_1 值没有达到 95% 也不意味着数据没有使用价值。相反, 这些数据仍然能够为我们提供丰富的信息, 可以根据具体场景来提升规范问题的判断质量。

(七) 自动标注与数据输出

训练出可用模型后, 就可以应用模型对未进行人工标注的裁判文书进行自动化标注。裁判文书库可以随着时间的推移、新文书的公开进行动态、实时补充。标注完成后, 便能将自动标注数据输出, 并与人工标注数据合并, 构成完整的数据集, 用于接下来的研究。除法实证研究外, 这些数据对判决预测、类案检索等法律人工智能应用场景也具有重要意义。

四、回到问题: 裁判文书数据在法学研究中的运用

正如本文引言部分的图 1 所示, 法实证研究需要从法学规范问题出发, 利用事实判断问题在法学规范问题和实证材料之间建立联系, 最终再经由事实判断问题的中介回到法学规范问题。研究开始时确立的法学问题意识不仅为研究设立了整体目标, 即回答法学规范问题, 也为数据分析

(87) 邓矜婷:《计算法学方法初阶》, 法律出版社 2021 年版, 第 194 页。

设定了目标，即回答事实判断问题。而法学问题和概念的拆解过程已经将比较宏观的事实判断问题拆解为了具体的问题并对应到了裁判文书中的要素，为裁判文书信息抽取和数据分析设定了框架。在通过信息抽取获得了裁判文书数据后，只须在对问题和数据的理解基础之上选取合适的统计分析方法^{〔88〕}，并将分析结论填入框架，从实证材料到解答法学规范问题的链条即可依次打通。本节将运用目前得到的数据，在分析“知假买假者起诉针对的商品多大程度上被法院认定为确实存在问题”的过程中讨论裁判文书数据的运用，同时对上文确立的问题意识进行回应。

（一）数据问题与法学问题的交织

在分析裁判文书数据时，对裁判文书数据的使用以及对统计结果的解释应当充分、谨慎地考虑裁判文书数据的局限性对结论的影响。在这一过程中，对数据的理解与对法学问题的理解交织在一起，只有既理解了法学问题，又理解了数据由何而来，才能正确运用数据分析法学问题。数据问题与法学问题的交织需要在具体的情景中讨论。

首先，惩罚性赔偿案件的裁判文书形成和公开时的不全面不会导致全部可得文书在内容上与全部案件产生结构性偏差，而裁判文书公开率的时间差异和文书公开的滞后性则使得文书数量的时间变化有迷惑性。图3呈现了2012年到2019年惩罚性赔偿案件（区分为知假买假案件和非知假买假案件）审结并公开文书的案件数量的季度变化，文书数量2014年之前极少，2014年到2017年迅速增长，2018年开始又大幅回落。裁判文书上网公开工作于2014年1月开始全面推行，到2015年6月底原则上实现了文书公开全覆盖。^{〔89〕}2014年之前的裁判文书基本是裁判文书网上线后各级人民法院补传的。2018年起惩罚性赔偿案件文书数的大幅回落也受到文书公开周期的影响。^{〔90〕}图3用到的裁判文书是在2020年1月检索收集的，当时还有很多已经生效的裁判文书还未来得及公开，会使得相应时间的文书数偏少。因此，相比于文书数量及其变化，不同类别文书的比例及其变化对真实情况的说明有更高的盖然性，也更具有讨论价值。

其次，在文书公开方面，惩罚性赔偿案件不是敏感案件，一般也不涉及法定不公开情形^{〔91〕}，虽然存在个别被告人反对公开或申请撤回的情况，但没有理由认为这足以对裁判文书数据产生结构性影响。至于裁判文书记载的内容，由于知假买假者的行为中能够影响裁判结果的部分应当被法官明确认定，故可以推定裁判文书的事实认定是对案件中知假买假者行为的客观体现。

最后，只有在案件中有规范评价意义的事实才会写入裁判文书，研究需要的信息有时不能在

〔88〕 关于如何利用数学统计方法进行事实描述和因果判断，从而检验假设或理论，在社会学等领域已经有十分成熟的讨论。这些统计方法并非本文讨论的重点，具体可参见〔美〕Donald J. Treiman（唐启明）：《量化数据分析：通过社会研究检验想法》，任强译，社会科学文献出版社2012年版；郭志刚主编：《社会统计分析方法：SPSS软件应用》（第2版），中国人民大学出版社2015年版。

〔89〕 从2014年1月1日起，全国中级人民法院都须上网公布裁判文书；同时，13个省（区、市）的基层人民法院也应当上网公布裁判文书。2015年6月底起，全国31个省（区、市）及新疆生产建设兵团的三级法院已全部实现生效裁判文书上网公布，即案件类型全覆盖、法院全覆盖。参见姜佩杉：《中国裁判文书网文书总量突破一亿》，载《人民法院报》2020年9月2日，第001版。

〔90〕 在2017年公开的全部裁判文书中，有39.4%的文书是在文书制作完成后超过90天才公开的。参见杨金晶、覃慧、何海波：《裁判文书上网公开的中国实践——进展、问题与完善》，载《中国法律评论》2019年第6期，第128-138页。

〔91〕 《最高人民法院关于人民法院在互联网公布裁判文书的规定》（法释〔2016〕19号），第4条。

裁判文书中直接体现。但是，裁判文书的优势之一便是蕴含的信息维度丰富，可以结合具体情况通过替代方案间接获取文书内容不能直接体现的信息。通过裁判文书数据分析知假买假行为，首先要面对的问题是在全部惩罚性赔偿案件文书中，哪些文书是涉及知假买假案件的？

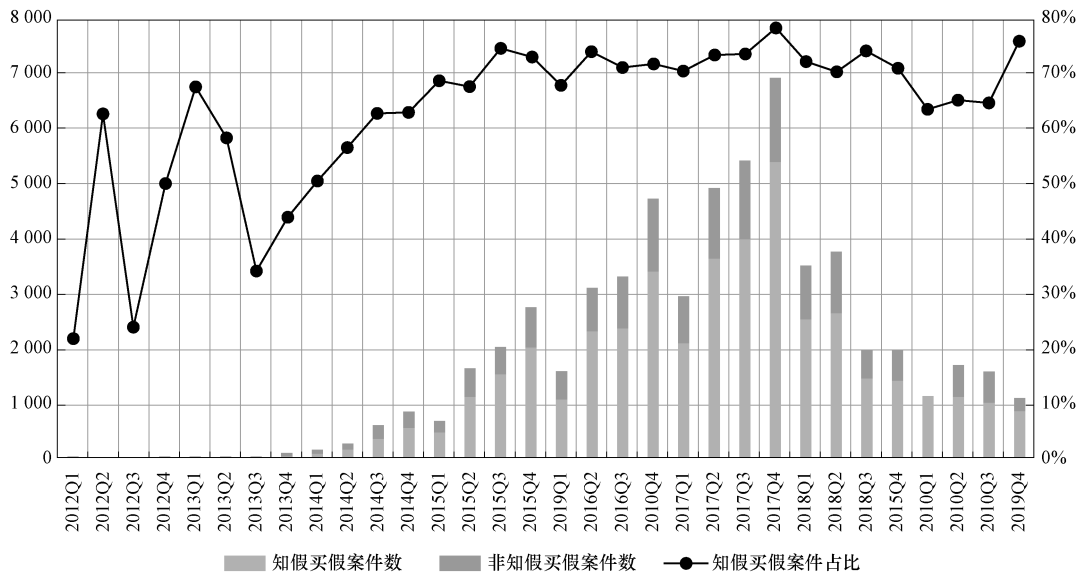


图3 2012年至2019年知假买假案件的数量与占比季度变化

裁判文书内容并不能很好地体现该案是否属于知假买假案件。在人工标注的文书中，法官明确在裁判文书中认定原告是否知假买假的仅占 15.8%。这并不意味着法官的裁判说理是错误的。因为对绝大部分法官来说，并没有查明原告是否是知假买假的必要。⁽⁹²⁾在认定知假买假的文书中，法官使用最多的认定依据是个案中原告行为有违正常消费习惯（占 48.4%）⁽⁹³⁾，其次是历史诉讼量（即原告提起消费者惩罚性赔偿诉讼的案件数，占 37.1%）。之所以历史诉讼量可以成为认定知假买假的依据，是因为普通消费者几乎不会频繁提起惩罚性赔偿诉讼，而知假买假者要想以惩罚性赔偿金作为收入来源就需要经常起诉获赔。以历史诉讼量作为判断依据对所有文书具有普遍适用性。不过，通过历史诉讼量判断原告是否是知假买假者得到的结果只是一种推定结果。例如，

(92) 第一，在因食品药品可能存在安全问题而起诉的案件中，不反对知假买假的法官可以通过援引《最高人民法院关于审理食品药品纠纷案件适用法律若干问题的规定》（法释〔2021〕17号）第3条的规定回避对知假买假问题的讨论。在所有法官未表示反对知假买假且打假事项为有毒有害的食品药品案件文书中，讨论了原告是否属于知假买假的仅占 4.2%。反对知假买假的法官更有认定是否是知假买假的必要，但在涉诉商品属于食品药品且打假事项为有毒有害的文书中，法官表示反对知假买假的仅占 10.9%。第二，在因欺诈起诉的案件中，不同法官可能会对《消费者权益保护法》第55条中“欺诈”的含义是应当依“商家单方的欺诈行为”来解释抑或是依民法一般规定来解释（即“消费者受欺诈产生认识错误，并基于认识错误作出意思表示”）存在差异。按照前一种方式解释《消费者权益保护法》中“欺诈”的法官没有必要关注原告是否明知商品或服务存在问题。只有按照后一种方式解释“欺诈”，反对知假买假的法官才有认定原告是否是知假买假的必要。但打假事项为“误导决策”的文书中，法官表示反对知假买假的仅占 13.6%。

(93) 有违正常消费习惯的购买行为，如短时间内多次购买同一或类似商品（重庆市沙坪坝区人民法院民事判决书（2017）渝0106民初19763号）、购买数量明显超出正常需要（江苏省泰州市中级人民法院民事判决书（2017）苏12民终2284号）、对购买过程全程录像（江苏省徐州市鼓楼区人民法院民事判决书（2016）苏0302民初3012）等。

历史诉讼量少但原告自认知假买假的也应当将案件认定为知假买假案件。目前，“知假买假案件”的判断标准是：一审提起诉讼量案数 ≥ 5 起的原告推定为知假买假者、知假买假者发起诉讼的案件视为知假买假案件。据此标准，裁判文书库中 75.34%的一审案件是知假买假案件。

（二）运用数据分析法学问题

运用裁判文书数据分析法学问题有两个层次：首先是通过裁判文书数据论证事实判断问题，然后通过事实判断问题的结论对法学规范问题进行结果论或目的论的论证。“知假买假者起诉针对的商品多大程度上被法院认定为确实存在问题”属于第一个层次的问题。

关于数据的解释利用，这里暂以“裁判情况\法官对事实的认定与审查\客观审查\审查结果\客观存在问题”要素为例，进行数据分析。该要素是指“经过法官审查，涉诉商品确实存在安全或欺诈问题”，相应模型的精确率和召回率如表 2 所示，其精确率可以达到 0.92，召回率 0.72 可以视为对裁判文书总体的 72%抽样。

表 2 本节问题用到的要素及其数据可靠性

要素名称	Precision	Recall
案件基本信息\文书情况\审判日期	1	1
裁判情况\法官对事实的认定与审查\客观审查\审查结果\客观存在问题	0.92	0.72

惩罚性赔偿裁判文书样本全库中共有截至 2019 年 12 月 31 日的 59 147 份裁判文书，其中审结日期在 2012 年至 2019 年的有 59 036 篇。如前所述，不同年份的裁判文书上网率存在不同，需要重点关注的是不同时间段的比例变化。2012 年到 2019 年每一季度的知假买假案件数量与占比如图 3 所示。2012 年和 2013 年的公开裁判文书数量过少，比例波动大。知假买假案件占比在 2012 年至 2015 年存在明显的上升趋势，此后大部分时间维持在 70%以上。这与 2014 年《最高人民法院关于审理食品药品纠纷案件适用法律若干问题的规定》第 3 条承认知假买假者的惩罚性赔偿请求权后相关制度没有大的变动相契合。总的来看，知假买假案件占惩罚性赔偿案件的绝大多数。

关于涉诉商品客观上是否存在安全或欺诈问题，如图 4 所示，在 2012 年至 2019 年的大部分时间，不管是否是知假买假案件，商品客观存在问题的比例都在 60%至 70%之间，总计比例分别为 61.7%和 62.3%。考虑“商品客观存在问题”的模型精确率和召回率，按照如下公式推算：

$$N = nP / R$$

（其中 N = 实际比例， n = 标注比例， P = Precision， R = Recall）

知假买假案件中和非知假买假案件中涉诉商品实际存在安全或欺诈问题的比例分别为 78.8%和 79.6%。

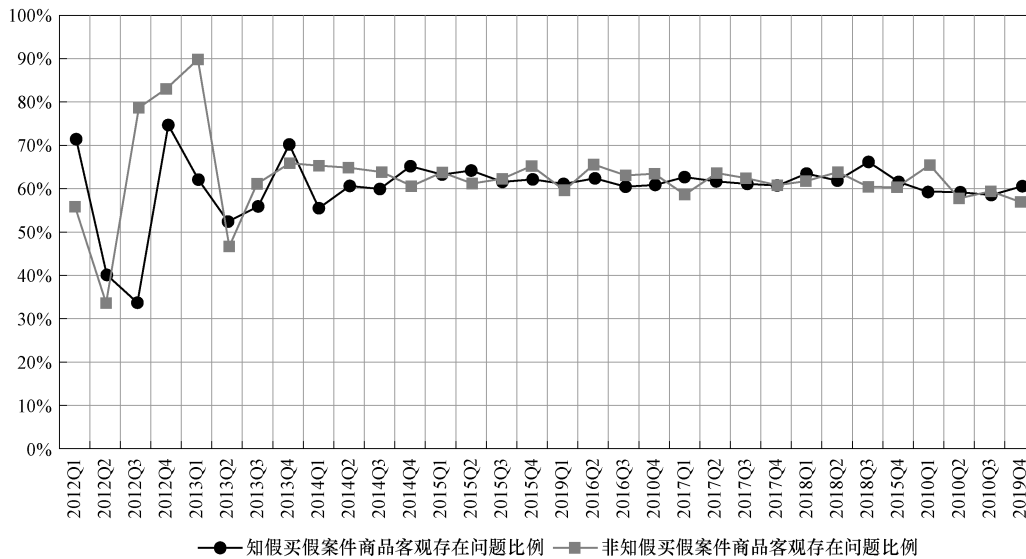


图4 2012年至2019年不同类型案件中商品客观存在问题的比例

上述数据回应了“知假买假者起诉针对的商品多大程度上被法院认定为确实存在问题”：在知假买假者起诉的接近80%的案件中，涉诉商品客观上确实存在安全或欺诈问题。这一比例与非知假买假者（即普通消费者）起诉时在对商品是否存在问题的判断上并不存在明显区别。即使个别知假买假者具有高度的投机性，也不足以在整体上影响对知假买假行为的评价。从这个角度看，允许知假买假者获得惩罚性赔偿金有利于惩罚和遏制不法行为、净化市场环境、保护消费者利益，整体上符合消费者惩罚性赔偿制度的立法目的。

最后，从事实判断问题回到法学规范问题，上述事实判断结论为“知假买假者应当获得惩罚性赔偿金”提供了一个目的论方面的支持理由。同时，知假买假行为的消极方面也不可忽视，对于知假买假行为的消极方面，还可以进行针对性的研究，从而为引导知假买假者的行为以“扬其利、去其弊”提供一个整体方案。例如，人民法院在承认知假买假者的消费者法律地位的同时，可以考虑适度提高对现行《中华人民共和国消费者权益保护法》第55条规定的“欺诈”和《中华人民共和国食品安全法》第148条规定的“不安全”的认定标准，让那些针对非实质性欺诈或者轻微安全瑕疵食品起诉索赔的知假买假者难以获利，从而在发挥这个群体正能量的同时，抑制其副作用。

当然，关于涉惩罚性赔偿裁判文书的信息抽取工作还可以不断优化，从而在细节层面不断提升前述结论的精准性。这也正是实证研究不断进步的体现。在此希望，本文提出的法实证研究思路和方法对后续类似研究具有一定的参考意义，也希望基于此种方法对涉知假买假者的惩罚性赔偿问题展开的后续研究对优化相关的学术认识和司法政策有所裨益。

The Empirical Legal Research Method Based on Judgement Data

—An Example from Intentional Purchasing of Defective Products

XIONG Bingwan WANG Junle

Abstract: Taking intentional purchasing of defective products as an example, this article comments on the empirical legal research method, through analyzing data extracted from judgments of punitive damages. Empirical legal research is a complete cycle, starting from normative jurisprudential problems, guided by teleology or consequentialism, deconstructing questions or concepts with problem orientation, theoretical resource and empirical evidence. It connects the legal problem with empirical information, and in turn analyzes facts better using empirical material, with the purpose of resolving relevant legal problems. Judgements are essential empirical materials. When conducting empirical researches with a large number of judgements, information can be extracted with precision from the original judgments via technologies like machine learning. At the same time, it should be noted that judgments are important yet limited representations of the legal practice activities, and it is necessary to work cautiously when using judgements for data analysis in empirical research, by assessing the validity of data rigorously and taking alternatives flexibly. Just like other field of scientific inquiry, empirical legal research does not offer a conclusive determination of the empirical world. It pursues an interpretation with higher probability and develops continuously with an open attitude.

Keywords: Empirical Research; Judgments; Information Extraction; Punitive Damages; Intentional Purchasing of Defective Products

(责任编辑: 许可 汪友年)