互联网企业刑事合规义务识别：分层、复合与技数赋能

谢　澍

摘要：涉案企业合规改革，本质上是为了鼓励更多企业建立健全的“事前”自主合规体系。企业事前自主合规体系建设的前提，在于有效识别刑事合规义务，这也是建构企业刑事合规管理体系的基础。互联网企业根据自身特点，在具体合规义务识别的过程中，应当坚持“分层”与“复合”的识别路径，确保“强制合规义务层级”与“优先合规义务层级”的分层识别，以及在此基础之上，行政合规义务与刑事合规义务、各类刑事合规义务来源的复合识别。互联网企业自主刑事合规体系需要“技数赋能”加以辅助，借助底层技术和数字建模相结合，探索生成式人工智能的介入，形成一套技术化、数字化、可视化的立体识别系统，将合规义务和风险评估嵌入企业日常业务流程，确保合规义务识别更加高效和准确。

关键词：企业合规；事前自主合规；合规义务识别；分层复合；技数赋能

中图分类号：D915.3　　**文献标识码：**A　　**文章编号：**1000—8691（2023）03—0131—08

一、事前自主合规与合规义务识别

“企业合规”是当前理论研究与实践探索中的热点问题，但随着最高人民检察院主导的涉案企业合规试点全面推行，研究者的关注重心似乎更多放置于“事中”和“事后”的企业刑事合规。尤其是以“合规不起诉”的兴起为聚焦点，甚至有观点将企业合规等同于一种涉案后的补救行为。“合规不起诉”是最典型的“事中”刑事合规，而企业的危机应对与配合执法也属于“事中”的范畴；而“事后”刑事合规，主要是指制定事后合规计划，以及计划的落实和复盘，确保刑事合规计划与合规监管协议发挥实质作用，推进企业刑事合规体系建设。有学者将这类企业在行政机关、司法机关的执法压力下，或在国际组织采取制裁措施的情况下，以减轻处罚或者取消制裁为目标所进行的合规举措，称为“合规整改”。[①] 然而，倘若将企业刑事合规仅仅局限于一种事中或事后行为，或许就偏离了涉案企业合规改革之根本目的。笔者认为，涉案企业合规改革，本质上是为了鼓励更多企业建立健全“事前”的自主合规体系，即日常性的、自发性的、前提合规关口的合规体系。唯有推动自主合规与被动合规的全面覆盖，以及事前合规与事中、事后合规的有序衔接，才能最大限度地从实体上阻断企业犯罪，实现企业治理的积极效果。近日，最高人民检察院印发的《关于加强新时代检察机关网络法治工作的意见》专门在“积极稳妥开展涉案互联网企业合规工作”中强调“探索以事后合规整改促进企业事前合规建设”，即遵循了上述思路。

企业事前自主合规体系建设的前提，在于有效识别刑事合规义务，这也是确立、制定、实施、评价、

基金项目：本文是中国犯罪学学会 2022 年度研究课题“涉互联网企业犯罪治理”（项目号：FZXXH2022C08）的阶段性成果。

作者简介：谢　澍，男，中国政法大学刑事司法学院副教授，主要从事刑事诉讼法学、证据法学研究。

① 陈瑞华：《有效合规管理的两种模式》，《法制与社会发展》2022 年第 1 期。

维护、改进企业刑事合规管理体系的基础。以国家市场监督管理总局、国家标准化管理委员会发布的《合规管理体系——要求及使用指南（标准号：GB/T 35770-2022）》（以下简称《合规指南》）为例，其中明确要求应当系统性地对合规义务进行识别[①]；同时，《合规指南》罗列的合规义务来源包括强制性遵守、自愿选择遵守和基于签署合同产生的义务[②]。然而，识别刑事合规义务的难点在于，刑事合规义务是随着立法、司法以及刑事政策不断变化的。例如，近年来，《中华人民共和国刑法》（以下简称《刑法》）增设了帮助信息网络犯罪活动罪、拒不履行信息网络安全管理义务罪、非法利用信息网络罪等罪名，督促企业配合相关行政监管履行责任义务；2022 年 4 月，最高人民检察院、公安部发布的《关于公安机关管辖的刑事案件立案追诉标准的规定（二）》，将企业常见的商业贿赂相关罪名入刑标准调低，对于标准的把握以及合规义务的内容呈现出动态变化的趋势。

对于近年来发展势头迅猛的互联网企业而言，由于企业在运行过程中需要处理海量的信息和数据，引发刑事合规风险的可能性也更高。数字革命创造出传统犯罪依然适应而新兴犯罪持续涌现的虚拟世界，犯罪样态也从传统的单一、集中转变为多元、分散。[③] 为了应对上述风险，《中华人民共和网络安全法》（以下简称《网络安全法》）、《中华人民共和数据安全法》（以下简称《数据安全法》）、《中华人民共和个人信息保护法》（以下简称《个人信息保护法》）以及《中华人民共和反电信网络诈骗法》（以下简称《反电信网络诈骗法》）等法律相继出台，持续强化着企业的安全主体责任。正是由于立法、司法以及政策向度的变化较为频繁，企业在面对新罪名、新规范、新标准时，可能陷入难以有效识别合规义务的困境。毕竟，互联网企业面对的刑事风险呈现多元化趋势，既有较为传统的商业贿赂、串通投标、侵犯商业秘密等风险，也有具备互联网特征的犯罪风险，如计算机犯罪、数据犯罪、创新业务非法经营、拒不履行信息网络安全管理义务等。可见，面对多元化的刑事风险，互联网企业试图界定较为清晰的合规义务，显然存在一定挑战。而随着 ChatGPT 等生成式人工智能工具的“出圈”，学界更应当思考，是否也可能在企业合规体系建设中探索类似的“技数”路径，辅助人们应对风险的多元化和复杂性。有鉴于此，本文试图结合笔者前期针对互联网企业自主刑事合规体系建设的调研成果，探索互联网企业刑事合规义务识别中的宏观理论指引与微观实践创新。需要说明的是，本文从互联网企业切入，但研究成果可能对于非互联网企业合规建设同样具有启发意义，只是互联网企业的相关问题更为突出，也更适应本文将要提倡的技数赋能之建构方向，因此为了集中论域，本文讨论的范围将聚焦于互联网企业的刑事合规义务识别。

二、互联网企业刑事合规义务识别的分层与复合

从企业自主合规走向企业被动合规，本是过往企业合规发展的历史进程。以域外为借镜，20 世纪中叶以前，自由主义思潮深度影响着美国，企业日常监管与外部监管之间缺乏法律规范的衔接，无法落实政府监管要求，为了强化企业监管、降低企业犯罪率，美国政府开始将合规计划引入法律实践，强制企业开展合规管理。[④] 然而，本文主张推动自主合规与被动合规的全面覆盖，以及事前合规与事中、事后合规的有序衔接，尤其强调企业事前自主合规体系建设，并不意味着逆转企业合规发展的历史进程。因为，当前语境下的企业事前自主合规体系，并非 20 世纪自由主义思潮下美国企业的自我管理。后者缺乏法律

① “组织应系统识别其活动、产品和服务所产生的合规义务，并评估其对组织运行所产生的影响。组织应建立过程以：a）识别新增及变更的合规义务，以保证持续合规；b）评估已识别的变更的义务所产生的影响，并对合规义务管理进行必要的调整。组织应保持其合规义务的文件化信息。”

② 具体包括，其一，组织强制性遵守的要求包括：法律法规；许可、执照或其他形式的授权；监管机构发布的命令、条例或指南；法院判决或行政决定；条约、公约和协议。其二，组织自愿选择遵守的要求包括：与社会团体或非政府组织签订的协议；与公共权力机构和客户签订的协议；组织的要求，如方针和程序；自愿的原则或规程；自愿性标志或环境承诺。其三，与组织签署合同产生的义务：相关组织的和产业的标准。

③ 谢澍、赵玮：《论网络犯罪案件的量刑证明——“整体主义”证明理论的实践探索》，《云南社会科学》2022 年第 1 期。

④ 万方：《合规计划作为预防性法律规则的规制逻辑与实践进路》，《政法论坛》2021 年第 6 期。

规范约束以至于无法落实监管要求，而前者建设的前提就在于有效识别刑事合规义务，前已述及，刑事合规义务的来源就包括需要强制性遵守的法律法规等。就此而言，从消极的企业自主合规，走向企业被动合规，再走向积极的企业自主合规，才是企业合规发展的应然路径。如果说事中、事后之涉案企业合规建设的有效性标准要求涉案企业在查明深层制度缺陷的基础上进行针对性制度修复，并且审查要点就在于涉案企业识别管控漏洞的准确性及深度[①]，那么，事前的企业自主合规就需要提前将可能的漏洞和缺陷进行识别，而漏洞和缺陷的标准即是企业合规义务所框定的，需要通过分层和复合的方法进行立体化形塑。当然，对于互联网企业而言，刑事合规义务的识别还要重视相关互联网刑事风险。

（一）刑事合规义务的分层识别

前已述及，《合规指南》中将企业合规的义务来源进行了罗列，包括组织强制性遵守的要求、组织自愿选择遵守的要求和与组织签署合同产生的义务。这实际上是对企业合规义务识别提出了“分层”的要求，即根据企业合规义务的强制程度进行界分，并有针对性地处理。对于互联网企业刑事合规义务识别而言，还需要准确把握互联网企业的显著特点及其所面临的互联网风险，尤其是关乎刑事风险的重点领域需要进入“强制合规义务层级”，否则无法有效预防刑事犯罪。

首先，互联网企业刑事合规义务识别的重中之重即是“强制合规义务层级”。这一层级主要包括来自刑事政策、法律法规、司法解释、指导性案例以及刑事判决。需要说明的是，刑事判决中的部分观点并不具有强制性，但在具体的司法实践中，尤其是同一辖区范围内的判决具有较高的参考价值，因此也应当被纳入“强制义务层级”。对于互联网企业而言，其强制合规义务不仅更新频繁，并且部分规范因为位阶不高而容易被遗漏。例如，互联网企业普遍开发了移动互联网应用程序，即我们常说的“手机App”，对此，2022 年 6 月 14 日，国家互联网信息办公室发布了修订后的《移动互联网应用程序信息服务管理规定》，其中就对相关合规义务进行了明确规范，包括应用程序提供者与应用程序分发平台应当共同履行的三项合规义务，以及应用程序提供者应当履行的十项合规义务和应用程序分发平台应当履行的五项合规义务。[②]可见，互联网企业的强制合规义务不仅涉及面广并且细碎，全面识别需要投入大量资源，仅仅依靠人工识别、手动识别难度较高。

其次，根据互联网企业的业务需求，在刑事合规义务识别的过程中应当区分“优先合规义务层级”。这一层级顺位处于“强制合规义务层级”之后，并非强制性要求，但因为与企业的业务开展有着密切关联，因此属于企业自愿遵守的合规义务。《合规指南》强调，组织应当将与业务相关的、最重要的合规义务加以识别。这也就意味着，“优先合规义务层级”中，可以根据与业务的相关程度以及业务本身的重要性进行再分层。《合规指南》提到，可以参考“帕累托原则”对此处合规义务分层进行指引。“帕累托原则”即“帕累托最优”（Pareto superior）[③]，即在对资源进行高效分配时，可能需要降低一部分人的效用，才能有效提升另一部分人的效用，而这种资源的配比往往呈现出“二八定理”，即 20% 的重要领域发挥着 80% 的积极效用，因此需要将资源尽可能地投入到 20% 的重要领域。当然，如果能在不减损一方福利的同时，通过资源配置的优化提升另一方福利，即是“帕累托改进”。[④]从企业合规的效率向度考察，合规效率并非合规成本与效益之间存在的简单比例关系，更不能一味地追求成本最小化，而是提倡对企业合规资源的有效整合与运用。刑事合规义务的识别，首先应当识别与核心业务最相关的部分，然后再

① 刘艳红：《涉案企业合规建设的有效性标准研究——以刑事涉案企业合规的犯罪预防为视角》，《东方法学》2022 年第 4 期。

② 王春晖、王巍：《App 提供者和分发平台的合规义务——解读〈移动互联网应用程序信息服务管理规定〉》，《中国电信业》2022 年第 7 期。

③ 当然，如波斯纳所言，在现实世界几乎不可能满足帕累托最优存在的条件，谈到效率概念，人们十有八九说的是卡尔多 - 希克斯效率（Kaldor - Hicks），相关论述可参见 [美] 理查德 • 波斯纳：《法律的经济分析》，蒋兆康译，北京：法律出版社，2012 年，第 16—17 页。

④ [印度] 阿马蒂亚 • 森：《伦理学与经济学》，王宇、王文玉译，北京：商务印书馆，2018 年，第 41 页。

逐步附带其他业务范畴。毕竟，企业的核心业务范畴可能仅占总业务的少量比例，但正是这部分核心业务能给企业带来大部分的收益，同时也面临更多的刑事风险，需要尽可能地通过企业合规义务识别将其纳入自主合规体系。以互联网企业必须面对的数据安全问题为例，《数据安全法》确立了“数据分类分级保护制度”，对企业而言，一方面，应当严格遵循国家重要数据目录和本地区、行业监管部门制定的重要数据具体目录对重要数据加强保护，这属于“强制合规义务层级”的范围；另一方面，对重要数据以外的其他数据，企业还需要自主进行分类分级并给予相应程度的保护，可以参考相关法律法规并根据所属行业的数据分类分级要求，结合企业自身实际情况，开展适合企业自身的数据分类分级保护举措，[①] 这就属于“优先合规义务层级”的范围。

此外，互联网企业刑事合规义务分层的意义还在于，区分不同合规义务的识别难度。法律法规由于最高度概括化的特点，可能存在晦涩难懂的情况，在刑事合规义务识别过程中相对更难被理解，即便存在相关司法解释，但也由于我国司法解释的“立法化”“抽象化”因素很难被理解。而刑事判决数量巨大，也最难实现全覆盖识别，更不必说在此基础上抽象出刑事判决中的各种观点，毕竟“同案同判”是理论界与实务界努力的方向，但尚未——实际上也不可能——完全实现，因而刑事判决的观点还可能存在各种冲突和矛盾。加之，同类刑事案件判决中的观点可能因为刑事政策或其他诸多因素而呈现出变化，尤其是在当下多项改革持续推进的时代，这种变化的速度相对较快，一旦没有及时、有效地进行追踪，企业刑事合规义务的识别就可能出现滞后。与“强制合规义务层级”相比，“优先合规义务层级”范围内的识别难度就相对较低，因为这一层级实质上是企业自愿遵守的合规义务，因而与企业本身的业务动态关联度较高，企业一般也能及时、准确、有效地加以把握。

（二）刑事合规义务的复合识别

在刑事合规义务分层识别的基础之上，还可以进一步优化刑事合规义务识别，进而发挥事半功倍的作用。《合规指南》强调，应当根据职能、部门、岗位、活动的区别，识别各职能、部门、岗位、活动中的合规风险源。这就需要根据企业各类业务进行企业合规义务的复合识别。申言之，从企业的业务运行过程考量，一个业务行为可能涉及多个刑事合规风险点，需要对相关合规义务进行复合识别，才能更好地落实合规管理。因为企业合规义务的复合识别本质上是服务于业务行为的合规，所以发起合规义务复合识别的出发点可以是基于识别各部门、职能和不同类型的组织活动中的合规风险源。就此而言，企业刑事合规义务的分层识别是基础性识别，应当是一个企业基于整体的合规需求而发起的，但企业刑事合规义务的复合识别则可以是企业内部各部门基于自身业务行为在分层识别基础上的“再识别”。

其一，行政合规义务与刑事合规义务的复合识别。对企业而言，刑事风险涉嫌的罪名大部分是行政犯，对互联网企业所涉及的互联网刑事风险更是如此。例如，拒不履行信息网络安全管理义务罪，即行政责任与刑事责任衔接的产物。而正如有学者指出的那样，该罪作为纯正不作为犯是义务犯，其实质根据在于对行为人所承担的社会角色和规范义务的违反，其不法内涵是通过特定的不履行积极行为义务表现出来的，因此，违反特定义务的人成为整个犯罪的核心角色和关键人物，其对特定义务的有意识违反奠定了正犯性。[②] 因此，预防拒不履行信息网络安全管理义务罪的前提即识别相关义务，包括行政法意义上的义务和刑事法意义上的义务。过往法律、行政法规规定的信息网络安全管理义务较为模糊和零散，互联网企业进行合规义务识别存在一定困难，但随着近年来涉及互联网犯罪和互联网监管的专项立法相

① 任文岱：《“数据分类分级保护”背景下的企业合规 以国家和公共利益为视角依法合规运用数据》，《民主与法制时报》2022 年 2 月 9 日。

② 就拒不履行信息网络安全管理义务罪而言，可能存在义务冲突的情形。这主要存在于该罪的网络服务提供者“致使违法信息大量传播”和“致使刑事犯罪证据灭失”之间。因为网络服务商要防止违法信息大量传播，最有效的方法是删除有关信息，但删除信息行为事后又有可能“致使刑事犯罪证据灭失，严重妨害司法机关依法追究犯罪”，这会令互联网企业无所适从，因此更需要在合规义务识别的过程中分析过往判决，把握义务冲突的解决策略。参见周光权：《拒不履行信息网络安全管理义务罪的司法适用》，《人民检察》2018 年第 9 期。

继出台，相关义务已较为清晰，除了前述《反电信网络诈骗法》的相关内容，《网络安全法》第 21 条也罗列了网络运营者应当按照网络安全等级保护制度的要求应履行的安全保护义务，《数据安全法》《个人信息保护法》对于相关义务亦有明确。可见，互联网企业在刑事合规义务识别的过程中，需要强调行政合规义务与刑事合规义务的复合识别，覆盖一般立法与专项立法及其司法解释中有关企业合规的义务性规定。

其二，各类刑事合规义务来源的复合识别。同一罪名可能涉及若干法律和司法解释，同时受到来自不同辖区范围内的诸多司法判决观点的影响，因此，需要在刑事合规义务分层识别的基础上，再对具体各类刑事合规义务来源进行复合识别。以互联网企业的数据安全合规为例，其风险来源包括但不限于侵犯国家数据安全、滥用数据垄断地位以及侵犯个人数据自决权等[①]，其规范依据既有《中华人民共和国民法典》《刑法》及相关司法解释，也有《网络安全法》《数据安全法》《个人信息保护法》《反电信网络诈骗法》等专项立法。更重要的是，相关风险的实践处理在不同地区存在一定差异，并且可能存在动态更新的趋势，因此需要互联网企业根据自身业务、地域对刑事合规义务进行复合识别，进而保障相关业务行为合规、有效地开展。

三、技数赋能互联网企业刑事合规义务识别

对互联网企业而言，针对刑事风险的分析评估可以分为两大部分：其一是法务发现风险并提起的个案评估，其二是重大事项的强制评估。但显而易见的是，无论是个案中的刑事风险还是重大事项的刑事风险，均是以企业合规义务为基本标准的，因此刑事合规义务识别也就成为互联网企业刑事风险之分析评估的首要前提。通常而言，企业风险的分析评估，包括但不限于风险甄别、分级分类、量化处理、方案制定、目标对齐等内容。上述分析评估举措指引并贯穿于后续的风险化解、效果核验以及优化改进等步骤之中，同时又与后续步骤形成闭环，呈现出全流程动态评估的样态。然而，仅仅将企业刑事合规义务识别与刑事风险分析评估进行衔接仍然是不够的。前已述及，互联网企业的合规义务不仅涉及面广并且细碎，全面识别需要投入大量资源，仅仅依靠人工识别、手动识别难度较高，即便进行了有效识别，仅仅依靠人力进行刑事风险的分析评估也可能存在效率低下且评估有效性存疑等问题。因此，互联网企业自主刑事合规体系尤其需要信息化建设加以辅助，其目的是适应互联网企业的显著特点，将合规义务和风险评估嵌入企业日常业务流程。具体而言，企业的信息化建设通常包括专门事项系统、一般项目系统、合同系统等，合规义务和风险评估的嵌入需要探索在系统卡点，并将刑事合规风险审查作为企业合规审查的关键内容，进而在系统内部实现刑事风险及时提示和合规事项长期留存，唯有如此，才能刚性地将刑事合规风险审查纳入到互联网企业制度建设层面。在此过程中，“技数赋能”的概念开始被部分互联网企业采纳并持续推进。有别于单一的“技术赋能”或“数字赋能”，“技数赋能”的优势即在于互联网企业将底层技术和数字建模相结合，形成一套技术化、数字化、可视化的立体识别系统，使得互联网企业刑事合规义务识别更加高效和准确，在进行业务操作的过程中可以一目了然地明确相关合规义务及可能的刑事风险，进而提前预防刑事犯罪。当然，需要明确的是，无论是大数据、人工智能，还是互联网企业刑事合规义务识别中所运用到的具体技数赋能举措，均是发挥辅助作用的，不能代替专业人才以及专业人才的法律知识和经验，否则可能异化为“技数主导”的局面。[②]

（一）技数赋能刑事合规义务识别的基本内容

互联网企业刑事合规义务的第一次识别是基础，但及时维护和更新刑事合规义务更为重要。《合规指南》强调，当企业内外之情形发生重大变化时，应当对合规风险进行周期性再评估。可见，企业应当

① 韩轶：《网络数据安全领域的企业刑事合规体系建构》，《江西社会科学》2023 年第 1 期。

② 谢澍：《人工智能如何“无偏见”地助力刑事司法——由“证据指引”转向“证明辅助”》，《法律科学（西北政法大学学报）》2020 年第 5 期。

对刑事合规义务进行常态化维护，以确保持续、有效的企业合规建设。立法和司法解释调整、司法政策变化、专项执法行动开展等因素都会导致有关企业合规的义务性规定发生改变，但企业合规义务的变化更多的还是体现在海量判决之中，很难仅仅依靠人工识别、手动识别加以完成，需要通过技术手段和数字建模实现完整提取和识别。具体而言，技数赋能刑事合规义务识别，需要建构可视化的立体识别系统，其中的技术基础即在于信息抽取，这是一种从无结构的自然文本中识别出实体、关系、事件等事实描述，以结构化的形式存储和利用的技术，目前已在部分互联网企业中开始运用。信息抽取的目标是，让机器理解互联网上的海量信息，作为知识图谱构建与填充、自动机器问答、信息检索、辅助决策等下游任务和应用的重要基础，为机器做正确决策提供大量相关知识。其中，基本方法包括：（1）实体识别，即从文本中识别出实体的边界和类别（来自预定义好的类别集合）；（2）关系抽取，致力于从文本中识别一对实体以及实体间的语义关系，构成关系三元组；（3）事件抽取，从文本中抽取出用户感兴趣的事件；（4）开放域抽取，直接使用句子原始字词片段作为实体之间的关系短语，而不是从固定的类型集合中选取短语，这也在一定程度上弥补了实体识别、关系抽取和事件抽取面向限定类别的知识抽取，难以应对未知域的问题。[①] 上述技术手段在互联网企业刑事合规义务识别过程中，重点针对宏观政策研究和微观操作指引、类型化指引提供支持。

首先，技数赋能刑事合规义务识别的宏观政策研究。刑事司法实践（判决）是刑事司法政策趋势的重要来源，通过宏观趋势的变动，能够更及时地了解政策方向；基于海量司法文书的标注和训练，实现文书中关键信息的抽取，可以帮助用户快速梳理案件事实并摘取所需信息，大幅度提升文书阅读效率。在技术探索层面，对大规模裁判文书中的司法要素识别是宏观数据分析的基础，区别于常规的实体识别问题，裁判文书的复杂长文结构增加了识别裁判文书中各要素（包括事实与证据、定罪与量刑等）之间的关联关系的技术难度，因此部分互联网企业在识别常规实体（人、事、地点）之外，增加了针对事实与证据、定罪与量刑之关联的关系抽取。[②] 此外，利用通用命名实体识别和预训练模型，识别公开裁判文书长文本中的案件事实、涉案金额、法律适用、判决结果等结构化要素，可以实现文书中关键信息的抽取。同时，利用 GRTE 模型结构，基于表格填充方法在复杂句子中抽取关系三元组，克服了仅关注局部特征而忽略了词语之间关系的全局把握，通过多次执行“生成—挖掘—集成”过程，可以逐步细化每个关系的表征，大大提升了司法要素之间关联实体的识别效果，为刑事合规义务识别打造大数据趋势分析。[③]

其次，技数赋能刑事合规义务识别的操作指引或类型化指引。以问答为具体形式的合规义务指引，能够增强可操作性。互联网企业自主刑事合规体系建设，需要从非结构化、复杂的法律中抽象出具体的合规要求，并采取合适的方式记录相关要求。业务操作指引在内容上应通俗易懂、简洁清晰，使其成为业务人员可随时查找、易于检索的工具。互联网企业需要将相关制度和技术文档以更便捷的形式给员工查阅，因此需要将刑事判决中的具体行为与法律、司法解释融合技术探索，基于强大的预训练生成语言模型，自动从各种形式长文档中挖掘问答对，为智能客服、智能问答等场景提供文档数据自动转化问答对的能力。并且，这种指引应当是类型化的，即在前述之“刑事合规义务的复合识别”基础上，根据不同的业务，给法务、业务人员制作类型化的操作指引。在技术探索层面，可以运用司法领域知识融合模型，构建司法知识图谱，并结合语言模型对段落级交互进行建模，实现段落和文档级别之单个实体的司法知识检索，为类案研究提供相似案件智能推荐的支持。[④]

① 郁博文：《图视角下的信息抽取技术研究》，安全内参官网，https://www.secrss.com/articles/49180，2023 年 2 月 21 日。

② Yanguang Chen, Yuanyuan Sun, Zhihao Yang, Hongfei Lin (2020). Joint Entity and Relation Extraction for Legal Documents with Legal Feature Enhancement. Proceedings of the 28th International Conference on Computational Linguistics, 1561–1571.

③ Ibid.

④ Yunqiu Shao, Jiaxin Mao, Yiqun Liu, Weizhi Ma, Ken Satoh, Min Zhang and Shaoping Ma (2020). BERT-PLI: Modeling Paragraph-Level Interactions for Legal Case Retrieval. Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20), 3501–3507.

（二）技数赋能刑事合规义务识别的样态呈现

借助互联网企业的信息化建设，以刑事合规义务为依据的系统卡点提示，可以尽可能避免企业人员因为自身专业知识不足或存在认知偏差[①]而遗漏合规义务。不同互联网企业在技数赋能刑事合规义务识别的探索中当然可以自主选择技术和方法，但其目标是总体一致的。

以互联网企业刑事合规中常见的涉爬虫类犯罪为例，通过对近年来涉爬虫类犯罪裁判的指控证据进行分析，不难发现，检方常见的指控证据包括：（1）鉴定意见，如鉴定意见认定涉案爬虫技术具有侵入性或破坏性等；（2）电子数据，如涉案数据、设备环境数据等；（3）证人证言、犯罪嫌疑人和被告人供述及辩解等。尤其是计算机视角下的危害计算机信息系统安全犯罪，鉴定意见具有绝对优势的证明力。因而，利用爬虫技术采集数据的数据属性、采集技术、采集目的是指控犯罪的证明重点，为了更好地帮助法务人员及时发现、中止涉爬虫类犯罪，在合规评审过程中，可以将风险行为进行分级评价：标记等级为“高”的被认定犯罪风险高；标记等级为“中”“低”的需要结合其他情况综合判断或认定为风险低或认定为风险可控。相应地，互联网企业可以在刑事合规义务分层与复合识别的基础上，根据采集对象、数据属性、技术手段和数据使用向度的差异进行具体的分值设置。分值设置主要可以从刑事风险兑现的两个层次考虑，首先是实质构罪，即直接满足构罪要件；其次是犯罪治理趋势变化，即一段时间内因为专项执法行动等因素，导致某类犯罪的判决显著增加，需要重点关注相关刑事风险。例如，近期开展的打击治理电信网络诈骗犯罪专项行动使得此类犯罪的判决增多，互联网企业就需要在合规义务识别的过程中重点关注此类犯罪风险。

具体而言，互联网企业可以从四个层面对涉爬虫犯罪的刑事风险进行评估，并对相关合规义务进行识别：（1）关于侵入政府等网站。其中，高风险行为包括破坏或侵入政府计算机信息系统涉嫌破坏计算机信息系统、侵入计算机信息系统罪等。（2）关于采集涉密信息、个人信息，以及采取技术保护措施的著作。其中，高风险行为首先包括采集国家秘密、商业秘密、个人信息分别涉嫌非法获取国家秘密罪、侵犯商业秘密罪、侵犯公民个人信息罪等；以及通过破坏或侵入技术手段采集受技术保护的著作涉嫌侵犯著作权罪等。而中风险行为有采集非公开其他数据，因为采集非公开其他数据通常伴随破坏或侵入技术手段。（3）关于采集手段。其中，高风险行为包括使用破坏或侵入技术手段，该行为涉嫌破坏计算机信息系统罪、非法获取计算机信息系统数据、非法控制计算机信息系统罪等。根据刑事司法实践，如绕过保护措施获取密码、绕过人机识别验证机制、SQL 注入、突破账号权限等，暴力破解 App、破解 URL 签名、破解 http 协议、反编译解析代码，以及影响对方服务器造成系统延迟、影响对方数据安全性、造成对方用户流失等。低风险行为包括采集频率干扰对方计算机信息系统正常运行，该行为涉嫌破坏计算机信息系统罪，根据过往刑事司法实践，这部分合规义务尤其提示关注采集对象为政府的场景。

上述应用场景的描绘，是以涉爬虫类犯罪为例，呈现技数赋能刑事合规义务识别的具体样态，涉及对立法、司法解释、宏观政策趋势、刑事司法实践（判决）的全面把握。当然，根据不同的合规义务，具体细节乃至风险评级都可以有所差异，但前提均是要建立在刑事合规义务识别的“分层”与“复合”之上，并借助技数赋能，形成一套技术化、数字化、可视化的立体识别系统，将合规义务和风险评估嵌入企业日常业务流程，确保合规义务识别更加高效和准确。

四、余　论

结合前期在部分互联网企业的调研访谈，本文初步梳理了互联网企业事前自主刑事合规体系建设中合规义务识别的必要性与可行性，这是一个企业合规理论与实务必须认真对待却又尚未深入展开研究的

① 过往针对认知偏差的研究主要集中于刑事诉讼程序开启后，但企业自主刑事合规在刑事诉讼程序开启前进行，其中认知偏差带来的影响也需要深入研究。参见谢澍：《从“认识论”到“认知论”——刑事诉讼法学研究之科学化走向》，《法制与社会发展》2021 年第 1 期。

领域。笔者认为，在具体合规义务识别的过程中，互联网企业应当坚持“分层”与“复合”的识别路径，确保“强制合规义务层级”与“优先合规义务层级”的分层识别，以及在此基础之上，探索行政合规义务与刑事合规义务、各类刑事合规义务来源的复合识别。当然，互联网企业刑事合规义务识别的方法革新，其目的是适应“技数赋能”的辅助路径。近期火遍各行各业的 ChatGPT，实际上就是建基于预训练生成语言模型的产物，那么，未来技数赋能企业合规是否也可能研发出类似的工具，自动从各种形式长文档中挖掘问答对进而提供有效辅助，即本文所憧憬的。需要强调的是，技数赋能互联网企业刑事合规义务识别，根本上仍是发挥辅助而非主导作用的，核心还是专业人才的法律知识和经验以及技术团队的技术化、数字化、可视化探索。互联网企业的核心竞争力在于掌握前沿知识和技术的专业团队，而互联网企业事前自主合规体系建设同样以此为基础，就此而言，企业合规的未来，即专业人才和技术人才的未来（在本文写作过程中还得到了胡图、赵楠的协助，在此谨致谢忱）。

Identification of Criminal Compliance Obligations of Internet Enterprises: Layering, Compounding and Digital Technology Empowerment

XIE Shu

(School of Criminal Law and Justice, China University of Political Science and Law, Beijing, 100088)

Abstract: The compliance reform of enterprises involved in legal cases essentially encourages more enterprises to establish and improve the independent compliance system in advance. The premise of constructing independent compliance system of enterprises in advance is to effectively identify the criminal compliance obligations, which is also the basis for establishing criminal compliance management system of enterprises. According to their own characteristics, Internet enterprises should adhere to “layering” and “compounding” identification path in the process of identifying specific compliance obligations, ensure the hierarchical identification of mandatory compliance obligation level and priority compliance obligation level, and on this basis, the composite identification of administrative compliance obligations and criminal compliance obligations, as well as the sources of various criminal compliance obligations. The independent criminal compliance system of Internet enterprises needs the assistance of digital technology empowerment. With the combination of underlying technology and digital modeling, a set of technical, digital and visual three-dimensional identification system is formed. The compliance obligations and risk assessment are embedded in the daily business processes of enterprises to ensure that the identification of compliance obligations is more efficient and accurate.

Keywords: Enterprise Compliance, Independent Compliance in Advance, Identification of Compliance Obligation, Layering and Compounding, Digital Technology Empowerment

[责任编辑：陈慧妮]