【知识产权】

新酒入旧瓶:企业数据保护的商业秘密路径

崔国斌

【摘　　要】商业秘密制度在企业数据产权保护中扮演着核心角色,却没有引起决策者和学术界的充分重视。企业收集的与经营活动有关的数据集合,通常都落入商业秘密法上的"经营信息"范围。即便数据条目为单纯的文学艺术作品,也不妨碍该数据集合整体上被视为企业的"经营信息",从而构成商业秘密法的保护客体。在判断数据集合的秘密性时,应区分数据条目与数据集合。数据条目来源于公共领域,并不妨碍数据集合整体上具有秘密性。网络平台通过前台向公众提供的数据集合条目,使得数据条目本身失去秘密性。不过,网络平台后台存储的受访问密码等有效保密措施控制的数据集合整体或其实质部分,依然满足"秘密性"的要求。公众破坏该后台保密措施直接获取该数据集合整体内容,将构成商业秘密侵权。沿着这一解释思路,大多数企业数据集合都能在商业秘密保护法的框架下得到有效保护。在商业秘密保护法之外,类似日本那样进行"限定提供数据"的平行立法,叠床架屋,没有必要;不区分公开数据与秘密数据的统一数据产权立法,也缺乏可行性。企业数据产权保护更合理的选择是回到现有的商业秘密保护法加可能的公开数据特殊保护立法的思路。

【关 键 词】数据产权;企业数据集合;商业秘密;秘密性

【作者简介】崔国斌,清华大学法学院教授、清华大学知识产权法研究中心主任(北京 100084)。

【原文出处】《政治与法律》(沪),2023.11.2~23

【基金项目】本文系国家社科基金重大项目"互联网交易制度与民事权利保护研究"(项目编号:20&ZD192)的阶段性成果。

一、引言

企业数据保护是近几年中国法学界的热点问题。笔者于本文中所称的企业数据是相对于个人信息和公共数据而言的宽泛概念,大体是指企业在生产经营过程中收集的源于自身或他人的具有一定规模的各类数据信息的集合,比如,电商平台收集的用户交易数据集合、社交媒体收集的用户个人信息集合、地图网站收集的地理信息集合、航空公司的航班信息集合、期刊网收集的论文电子版文档集合,等等。它与传统的商业秘密类信息(如技术方案、程序代码、客户名单、招投标定价信息、经营企划方案等技术或经营信息)相比,主要区别在于内容性质、产生过程和数据规模。我国企业数据集合产权保护领域的最新且最重要的政策进展是中共中央和国务院2022年底发布的《关于构建数据基础制度更好发挥数据要素作用的意见》(以下简称:《数据二十条》),其第三条提出所谓"三权分置"思路:"根据数据来源和数据生成特征,分别界定数据生产、流通、使用过程中各参与方享有的合法权利,建立数据资源持有权、数据加工使用权、数据产品经营权等分置的产权运行机制。"

从表面上看,商业秘密保护法[①]似乎是企业数据产权保护立法无法绕过的关键制度。首先,从现行法对于商业秘密的定义看,它覆盖范围极其宽泛,根据《中华人民共和国反不正当竞争法》(2019年修订,以下简称:《反不正当竞争法》)第九条第四款,商业秘密是指任何"不为公众所知悉、具有商业价值并经权利人采取相应保密措施的技术信息、经营信息等

商业信息”。企业数据大多为非公开数据,通常具有一定的商业价值,为保密措施所覆盖,因此很容易就落入所谓“商业信息”的字面意思所指涉的范围。2020年《最高人民法院关于审理侵犯商业秘密民事案件适用法律若干问题的规定》(以下简称:《侵犯商业秘密规定》)第一条也已经确认,企业数据可以构成商业秘密。因此,在企业数据产权保护中,商业秘密保护法很可能要扮演重要甚至核心角色。[②]沿着这一思路,《数据二十条》的“三权分置”结构就会被重新解构:它所列举的数据资源或数据产品,在公开之前,很可能构成企业的商业秘密;“数据资源持有权”则对应于企业对自己生产或收集的数据所享有的商业秘密权;[③]“数据加工使用权”对应于数据加工者从数据生产或收集者那里获得的商业秘密许可使用权;“数据产品经营权”应该是指数据加工者在商业秘密许可范围内生产数据衍生产品的权利以及对衍生产品重新主张的商业秘密权或其他知识产权。[④]在此基础上,《数据二十条》第七条希望健全的“数据相关财产性权益”的流转机制,自然就对应于商业秘密的转让和许可机制。显然,在这一思路下,绝大部分秘密的企业数据都被商业秘密法有效覆盖,不再需要统一的数据产权立法。剩下的是一小部分公开的企业数据,可以通过《反不正当竞争法》的原则条款获得保护,或者将来通过专门立法获得更精细的保护。

不过,遗憾的是,《数据二十条》规划的体系并没有商业秘密保护法的位置。该文件虽然在“总体要求”部分提到数据产权制度建设以“保护个人信息和商业秘密为前提”,但是在后面的具体制度建议中再也没有提及商业秘密。因此,在《数据二十条》所描绘的“三权分置”的框架下,现有的商业秘密保护法扮演何种角色,并不清楚。这给未来数据产权制度建设带来巨大的挑战:如果大部分企业数据构成商业秘密,那还需要统一的数据产权立法吗?如果一定要有统一立法,那会取代商业秘密保护法吗?如果不取代商业秘密保护法,那么拟议的统一立法如何与商业秘密保护法协调?

上述分析表明,如何看待企业数据商业秘密保护的可能性,对决策者后续落实《数据二十条》,完善我国数据产权保护制度的思路会有重要影响。如果决策者承认商业秘密保护法的关键作用,数据产权立法的首要工作就将是完善现有商业秘密保护法,附带考虑次要的公开数据保护立法的问题。在现有制度可堪大用的情况下,立法者并没有必要叠床架屋去引入全新的统一立法,同时又制造出更多的、更复杂的新旧制度协调问题。

依据现有主要的我国商业秘密保护法即《反不正当竞争法》,一项信息要满足以下条件才构成商业秘密:(1)构成所谓“商业信息”(客体审查要件);[⑤](2)具有秘密性(不为公众所知悉);(3)具有商业价值;(4)权利人采取了合理的保密措施。表面看来,将这些要件套用至企业数据集合,似乎顺理成章,但实际上并非如此简单。源于机械时代的商业秘密制度所保护的“商业信息”,过去更多的是产品配方、工艺流程、客户信息、合同报价之类的传统商业秘密的客体。这些大多是企业从无到有创造出来的、体现个性化劳动、规模较小的数据信息。网络时代的企业收集完成数据集合,其数据常常来源于第三方或公共领域,处于原始未经深度加工的状态,也未体现企业的个性化劳动,具有远超传统商业秘密的信息规模,等等。两相对照,非传统的“企业数据集合”是否应当被视为商业秘密保护法意义上的“商业信息”类型,是否具有秘密性,很容易受到质疑。比如,常见的期刊论文数据库、法律法规案例数据库、社交媒体用户生成内容数据集合,其数据条目既非技术信息,也非传统的经营信息(客户名单、合同报价与商业计划书等),而是通常不受商业秘密法保护的文学艺术作品,这些作品数字化文档的“集合”是否属于商业秘密法意图保护的“经营信息”类型?很多企业数据集合整体上不对外公开,但是其中的数据条目常常源自公开渠道或处在公开状态,这是否妨碍该数据集合整体的“秘密性”?企业在收集数据时付出了实质性的劳动或投资,但并未付出个性化或创造性劳动,这是否会影响该数据集合的客体属性或“秘密性”?

对于这些问题,国内很多司法判决和学术研究成果都给出否定性答案。[⑥]这导致很多企业以为无

法利用商业秘密法来保护自己的数据集合，转而寻求《反不正当竞争法》的原则条款的保护。本文第二部分从商业秘密保护法下经营信息的多样性入手，突破传统认识误区，分析将典型的企业数据集合归入“经营信息”类别的合理性。本文第三部分关注商业秘密的秘密性要件，分别探讨了“数据条目来源于公共领域”“收集工作体现了实质投入但无创造性”“整体保密但部分条目对外提供”等三种典型企业数据集合的秘密性。与社会上普遍存在的误解不同，笔者认为它们依然具备商业秘密的“秘密性”。本文第四部分对商业秘密保护之外的替代性立法思路做出回应，认为日韩式的“限定提供数据”平行立法，制造的问题比解决的问题还多，不值得效仿；不区分公开与秘密数据的统一数据产权立法，无法真正取代商业秘密保护机制。笔者最后的结论是，沿用现有的商业秘密保护机制就能够有效应对数据产权保护的主要挑战，没有必要进行颠覆性的法律制度变革。

二、企业数据集合的经营信息属性

在传统商业秘密法的框架下，一项信息是否落入受保护“经营信息”的范围，较少受到关注。过去，关于商业秘密客体审查的学术讨论主要集中在具有技术和作品双重属性的计算机程序上，即它作为版权保护客体的同时，是否也可以成为商业秘密法的保护客体。[7]除此之外，美国还有关于剧本创意是否落入商业秘密法下“商业信息”范围的学术争议。[8]现在，企业数据集合可能是新的需要认真考虑其商业秘密客体属性的对象。

(一)数据集合内容的多样性

从美国法的历史看，商业秘密法的保护客体最初主要是技术方案，[9]然后逐步扩展到各种形式的商业信息。[10]现在，从商业秘密定义和国内外立法、司法解释的示例看，商业秘密客体的保护范围很广。[11]比如，2016年，美国联邦层级的《保护商业秘密法》将商业秘密定义为任何形式和类型的财务、商业、科学、技术、经济或工程信息。[12]美国有学者感叹，现在不经过复杂诉讼，被告很难证明企业所持有的任何信息不构成商业秘密。[13]中国现行法所确定的商业秘密范围也很宽泛，从2019年修订的《反不正当竞争法》第九条第四款的字面看，商业秘密涵盖任何具有商业价值的商业信息(技术信息和经营信息)。其中，技术信息的含义比较明确，经营信息的含义则相对模糊。依据2020年《侵犯商业秘密规定》第一条第二款，经营信息包括“与经营活动有关的创意、管理、销售、财务、计划、样本、招投标材料、客户信息、数据等信息”。这里提到了“数据”，但没有明确其具体含义。最高人民法院的参与起草者后来解读这一司法解释的文章也没有将这一点作为要点进行解释。[14]

企业收集的数据集合是否落入商业秘密客体的范围，与数据条目的内容有直接的关系。如果数据条目是《侵犯商业秘密规定》第一条第一款所称的“技术信息”，则数据集合无论是划入“技术信息”或“经营信息”类别，都可以获得商业秘密法的保护，对此应该没有什么争议。如果数据条目并非技术信息，则大多落入“经营信息”范围，也可以轻松纳入商业秘密的保护范围。比如，网络平台收集的用户注册信息集合，其数据条目包含用户ID、密码、头像、教育经历、联系方法等内容。此类数据条目与传统的客户信息类商业秘密可以直接类比，构成商业秘密，应该不成问题。[15]新浪诉脉脉案的诉争数据集合是这一方面的典型代表。[16]再如，电商平台、证券机构收集的用户交易数据，医疗机构积累的病人医疗记录数据，[17]产品售后的用户信息反馈数据等，也都直接与司法解释中所列举的“销售”“财务”和“管理”经营活动相关，显然也是典型的商业秘密保护客体。淘宝诉美景案是涉及此类数据的典型案例。此外，企业物联网收集的设备运行数据、地图导航数据、公共车的时刻信息，也都落入“经营信息”的范围。深圳公交公司案是典型案例，但法院并没有按照商业秘密的思路来处理，十分遗憾。[18]

虽然商业秘密法列举的“经营信息”极其宽泛，覆盖了绝大多数企业的数据集合，但是还是留下了一项重要的空白——文学艺术作品。从中外的立法或司法解释看，文学艺术作品从未被商业秘密立法作为可能的保护客体直接加以列举。比如，美国的商业秘密联邦立法虽然不厌其烦地罗列了多种商业秘密的客体，却并未提及普遍存在的文学艺术作

品。在中国,最高人民法院的商业秘密司法所详细列举的典型“经营信息”示例中,尽管很多也可能构成作品,比如记录经营创意、管理方法、销售计划、财务报表的文字作品、招投标文字作品、呈现样本的美术或摄影作品等,但法院并没有直接提到“文学艺术作品”。最高人民法院应该是认为,内容与经营活动无关的单纯文学艺术作品,比如小说、剧本、视听作品、书法或绘画等作品本身,并不构成这里所述的“经营信息”,不能以商业秘密获得保护。

既然文学艺术作品被排除出商业秘密客体的范围,那么由文学艺术作品条目组成的数据集合,是否也因此被排除出商业秘密法的保护范围呢?前面提到的期刊网论文数据库、法规案例数据库、社交媒体用户内容集合、图片库中摄影作品集合、大众点评类用户评价数据集合,[19]就是这方面的典型例子。显然,这一类数据集合很有代表性,并具有重要的经济价值。企业耗费实质成本收集了这样的作品集合后,很可能会采取保密措施阻止竞争对手接触和利用该数据集合。如果竞争对手刻意规避保密措施,盗取并利用该数据集合,则会引发商业秘密保护法是否适用的问题。因此,我们有必要深入探讨此类重要的数据集合的商业秘密客体属性。

(二)排斥文学艺术作品的原因

中外商业秘密法排斥文学艺术作品,表面的原因是,单纯文学艺术作品并不属于与经营活动有关的信息。不过,这一解释并未触及底层的政策性考虑。毕竟,利用商业秘密法保护未发表的文学艺术作品,原本并不产生操作层面的困难。比如,计算机程序代码就获得著作权法和商业秘密法的双重保护。此外,“与经营活动相关”的信息,原本就有很大的解释空间。对于专业的创作者或制片公司而言,文学艺术工作者创作文学艺术作品与软件开发者编写软件代码并无本质差别,很难说与经营活动无关。法律上厚此薄彼,需要更底层的政策性解释。在笔者看来,底层的原因可能有多个方面。

其一,在著作权法自动保护文学艺术作品的背景下,作品发表前提供商业秘密的重叠保护通常是多余的。对于作品中的独创性表达,著作权法禁止他人未经许可对它的发表、复制、演绎和传播,提供了充分的激励机制。潜在的商业秘密保护虽然并不完全与著作权保护重叠,但没有实质超出这一范围。[20]对于作品中记载的不受版权保护的技术方案、实验数据、地理信息、客户信息(客户名单)、交易信息或经营计划(思想)等内容,如果构成“技术信息”或“经营信息”,则依然能够获得商业秘密法的补充保护。当然,前提是它们还要符合商业秘密保护的秘密性、保密性和价值性等要件。

由此看来,文学艺术作品中真正处在空白地带,既不受著作权法也不受商业秘密法保护的内容,其实相当有限,可能仅仅是非常抽象的作品创意、艺术风格、故事情节之类的内容。它们既非商业秘密法上的“经营信息”,又非著作权法意义上的“表达”。即便立法者愿意对这部分抽象思想提供商业秘密法的补充保护,也很可能会因为这些内容本身过于抽象或边界过于模糊,难以确定客体边界,而无法在操作层面落实该保护。[21]比如,在具体的侵权个案中,如果抄袭的部分过于抽象,被控侵权者是否侵害了权利人的“商业秘密”就很难判断。因此,在文学艺术作品的著作权保护的基础上,提供商业秘密的重叠保护,没有太大的意义。

其二,商业秘密法原本就无法有效保护文学艺术作品的表达。商业秘密法只能在作品发表前提供有限的保护,而大部分文学艺术作品要实现自身的商业价值,都要对外公开发表并传播。这与商业秘密法的保密要求天然矛盾——一旦作品公开发表,商业秘密保护就失去用武之地。因此,没有必要将文学艺术作品视为典型的商业秘密法保护客体。

其三,商业秘密法对单纯文学艺术作品提供重叠保护,可能会在一定程度上影响著作权法激励机制的运作。对文学艺术作品提供商业秘密的替代性保护,有可能降低部分著作权人公开发表作品的意愿,转而更多地选择利用商业秘密许可机制控制作品的传播,从而将保护延伸至事实消息、技术方案、实验数据等不受版权保护的客体内容,也可能延伸至著作权法原本并不限制的功能性使用行为。正是基于商业秘密保护可能妨碍版权法立法目的这一原

因,美国联邦版权法就不许可权利人依据州法对受联邦版权法保护的内容提出平行的商业秘密保护主张,此即所谓的先占或排斥(Preemption)学说。[22]法院判断商业秘密保护主张是否被联邦版权法先占时,考虑下面两项要件:主张保护的客体是否落入版权客体的范围;所要保护的权利是否等同于版权法赋予的排他权。[23]

上述前两项理由只是说明,商业秘密的重叠保护没有太大意义,但也没有实质危害;而第三项理由涉及重叠保护对著作权激励机制的负面影响,值得更认真地对待。不过,这种负面影响同样很有限。首先,与技术方案不同,文学艺术作品的表达有充分的可替代性,部分作者选择商业秘密路径保护作品,对社会的负面影响微乎其微。其次,商业秘密法对未发表的文学艺术作品提供补充保护,不影响公众在公共领域的行动自由。最后,商业秘密保护对于著作权人发表作品的积极性的影响也几乎可以忽略不计。过去,人们对商业秘密法保护技术秘密是否会影响发明人申请专利积极性存在类似的担心。美国联邦最高法院认为,技术方案的商业秘密保护并不实质影响专利法激励披露的公共政策,[24]这也是世界各国立法者的共识——否则就不会出现商业秘密法与专利法在各国共存的局面。著作权法原本就有禁止未经许可公开发表他人作品的保护机制,同时,绝大部分作品的著作权人通过作品的公开传播才能获得实质性的回报。在此基础上,有理由相信,即便单纯文学艺术作品获得商业秘密的重叠保护,也不会对著作权人发表作品的积极性产生实质影响。

正因为商业秘密的重叠保护没有实质性的危害,美国法院在适用联邦版权法的先占规则时,并不绝对排除各州商业秘密法的重叠保护。如果被告除了接触和利用相关版权客体内容,还额外地违反了保密义务,则州法关于商业秘密保护的规则不被排除。[25]这大大限制了联邦版权先占规则的适用范围,使很多州法上的商业秘密的补充保护主张得到支持。显然,法院更看重商业背信行为的可谴责性,从而更多地选择无视重叠保护的微不足道的负面影响。此外,值得一提的是,自从美国2016年通过联邦层级的《商业秘密保护法》(Defend Trade Secrets Act)之后,依据该法提出的商业秘密主张就不再被联邦版权法先占或排除了,因为《商业秘密保护法》本身也是联邦法律,不存在被联邦法先占的问题。[26]因此,美国法下版权与商业秘密的重叠保护的可能性进一步增加。

综上,商业秘密法排斥文学艺术作品,实现的公共政策收益有限。许可重叠保护,负面影响也同样有限。在可预见的未来,这一结论都不会有太大的变化。因此,社会没有放弃或改革这一传统的迫切需要,商业秘密法很可能会继续延续其排斥文学艺术作品的习惯做法。

(三)数据条目与数据集合客体属性的区分

如前所述,商业秘密法排斥文学艺术作品表达,并非出于重要的著作权法公共政策的考虑。即便商业秘密法重叠保护作品表达,也不会带来实质的负面影响。在此基础上,商业秘密法保护由众多作品条目组成的数据集合,损害著作权法公共政策的可能性就更小。原因很简单,多数情况下,数据集合只有其中作品条目达到一定数量并且整体上具有秘密性之后,才能作为商业秘密客体获得保护;该保护也仅仅限制公众获取和利用数据集合整体或实质部分,并不延及单个或有限数量的作品条目。因此,公众对数据集合中公开的单个或有限数量的作品条目,及其中不受保护的事实或思想的自由利用,通常不受数据集合的商业秘密保护的影响。

当然,将文学艺术作品的集合视为商业秘密保护客体,负面影响很小,这还只能说明决策者可以对它提供商业秘密保护。更重要的问题是,是否有必要提供此类保护?答案是肯定的。当数据条目与数据集合在观念上可以相互区分时,通常意味着二者体现了不同主体的劳动、资本投入或人格利益,法律需要分别向二者的创作者或投资人提供各自的保护路径。比如,如果企业数据集合的数据条目是用户或第三方创作的文学艺术作品,该用户或第三方可以通过数据条目的著作权来保护自己的投入,而企业作为数据条目的收集者,并不能直接依据数据条目的著作权来保护自己的投入,因此需要平行的制

度安排来保护该投入。商业秘密保护刚好可以填补这一制度空白。在这一保护机制下,数据集合的收集者能且仅能阻止别人通过不当手段从它那里获取数据集合本身,而不能阻止别人对单个或有限数量的不满足秘密性要求的数据条目的利用。因此,商业秘密法保护数据集合,仅仅保护收集者在数据收集过程中的劳动或投入,而不涉及该集合中的单个或有限数量的作品本身所耗费的劳动和投入。[27]后者由著作权法或其他相关法律提供保护,与商业秘密法无关。

其实,在观念上区分作品表达与其中包含的商业秘密,分别提供著作权和商业秘密权保护,已经是很成熟的做法。如前所述,商业秘密法所保护的"经营信息",很多都蕴含于作品(含单纯的文学艺术作品)之中。比如,商业秘密法所保护的经营创意、管理方法、销售计划、财务报表、招标策略、产品样品、地理信息等"经营信息",大多记录在文字作品、美术作品、摄影作品、卫星照片甚至是视听作品中。记录这些经营信息的媒介呈现出文学艺术作品的外观,但这并不妨碍商业秘密法对它们所记录的"经营信息"内容进行保护。沿着同样的思路可知:文学艺术作品条目虽非商业秘密保护客体,但这不妨碍网络平台为经营目的收集的含有作品条目的信息集合在整体上构成商业秘密法的保护客体。

在处理由个人信息的数据条目组成的数据集合的商业秘密客体属性时,法院实际上已经习惯这一思路。个人信息条目并非商业秘密客体,但这并不妨碍网络平台收集的个人信息集合整体上构成商业秘密客体。[28]含有作品条目的数据集合与含有用户交易信息或用户个人信息的数据集合,从信息汇编的角度看,表现形式并无本质区别;收集者投入的劳动的性质和数量,并无明显的区别;两类数据集合的应用场景和所要满足的经营目的,也无本质差别。因此,无论是出于保护投资、制裁背信行为,还是为了避免自助措施导致的社会资源浪费,[29]商业秘密法都没有明显的理由区分含有作品条目的数据集合与含有用户交易信息或用户个人信息的数据集合,拒绝保护前者却保护后者。

在观念上接受数据条目与数据集合二分的思路后,如何处理数据收集者的商业秘密权和数据条目所体现的在先权利人的利益的冲突,就成了数据商业秘密权属方面的重要问题。这里所说的数据条目所体现的在先权益,可能包含用户的隐私、个人信息、肖像、作品、商业秘密等。在现有的关于数据产权的讨论中,很多学者因为数据条目中在先权益的存在,而无法接受或理解数据收集者对数据集合享有商业秘密权的观念。比如,在数据条目的内容为个人信息时,很多学者可能就无法理解,为什么在无数用户的个人信息权利之上,数据收集者可以对用户个人信息的集合享有数据财产权(商业秘密权)。[30]

其实,接受数据条目与数据集合相区分的观念后,从法律上协调数据收集者与数据条目所涉第三方的利益关系变得很简单。数据集合商业秘密权像传统的知识产权一样,也是一种消极权利,[31]仅能够保证权利人对自己收集并控制的秘密信息的控制,即排除他人未经许可的接触、使用和披露该数据集合,并不能保证权利人自己可以主动使用和披露相关的数据条目信息。如果数据条目涉及第三方在先权益,则数据收集者利用和传播数据集合的内容,需要获得第三方的授权。比如,数据本身体现了自然人的个人信息,数据收集者在收集、处理和利用个人信息时,应当依据《中华人民共和国个人信息保护法》获得授权。法律在处理数据收集者的商业秘密权与他人隐私、著作权、国家秘密等冲突时,基本思路是一致的。总之,数据收集者对数据集合整体所享有的商业秘密权与数据条目中第三方的在先权益,在观念上相互独立,相互牵制,并非单纯地相互排斥。

三、企业数据集合的秘密性

秘密性是商业秘密获得保护的最重要的要件之一。依据《侵犯商业秘密规定》第三条,秘密性是指有关信息"不为所属领域的相关人员普遍知悉和容易获得"。这里对秘密性作出了双重的要求,其一,它事实上处于相对秘密状态("不普遍知悉");其二,"不容易获得",即从其他渠道"获得该项信息要有一定的难度"。[32]其中,"不普遍知悉"要求,确保公众在

公共领域的行动自由不受影响;“不容易获得”要求则进一步提高门槛,只有不当获取商业秘密通过合法途径难以获得因而能够给权利人带来实质性的竞争优势时,该商业秘密才能获得保护。依照这一双重标准,如果公众无需实质投入(无需付出一定的代价)就能够通过第三方公开渠道或反向工程而获得该信息,则该信息处在“容易获得”状态,[33]不具备秘密性。

将上述秘密性标准应用于企业数据集合,可能存在以下几方面的疑问:其一,很多数据集合的条目原本处在公开状态,这是否会影响数据集合本身的秘密性?其二,企业虽然有实质投入,但并未作出创造性贡献,这一事实是否足以认定该数据集合“不容易获得”,因而具有秘密性?其三,网络平台对外提供服务,许可用户获取数据条目,但对数据集合整体加密,这是否影响数据集合的秘密性?

(一)数据条目与数据集合秘密性的区分

与客体审查环节类似,在判断数据集合的秘密性时,我们也要区分数据条目与数据集合本身的秘密性。如果数据条目具有秘密性,则通常包含该数据条目的数据集合本身也具有秘密性。“比如,物联网上私人设备产生的单个用户数据、医院的个人病历数据、电子商务平台后台生成的单个用户的交易数据、物流和航空公司用户的个人行程数据等,都因为收集者采取了保密措施而使公众无法通过公开渠道获得。数据收集者之间的数据交换通常都是通过保密渠道进行的。单个用户的数据具有秘密性,这些数据的集合自然也具有秘密性。”[34]

值得一提的是,关于私人设备收集信息的秘密性,德国的德莱克瑟(Drexl)教授似乎有不同的意见。他认为工厂内的机器产生的信息有秘密性,而在开放道路上行驶的汽车所收集的信息不具有秘密性,因为其他汽车厂商也可以自由收集相同的信息。[35]国内也有学者认为,公共场所的传感器收集的数据不具有秘密性。[36]其实,这里的关键不是其他产商可不可以在相同场所自由收集——商业秘密法原本就不禁止他人独立收集或对公开渠道的产品进行反向工程。我们关注的不是街道场景的开放性,而是记录在载体上的数据集合本身的秘密性。[37]如果数据条目的收集需要实质投入,则意味着“不容易获得”,从而满足秘密性要求。不只如此,如果企业对其中原本就具有秘密性的数据条目进一步选择、加工和编排,则会进一步增加数据集合本身的秘密性。比如,医疗机构通常会对医疗数据进行加工并作匿名化处理,使得数据集合中数据条目信息的存在状态与原始数据的形态有很大差别,这些新形态的数据条目的结合就更可以被视为秘密信息了。[38]

实践中,大多数企业的数据集合从数据条目到数据集合整体都不曾对外公开,或者部分数据条目公开而部分条目不公开。依据上述标准,这些数据集合大多应该能轻松满足秘密性要求。[39]真正引发争议并且需要认真对待的是,所有数据条目通过前台公开可得,而数据集合整体存储在后台,并不对外公开的情形。国外有学者认为,即便付出了很多时间、金钱和精力对公开的信息进行汇编,该汇编结果也不构成商业秘密。在这一意见的秉持者看来,即便该汇编结果具备了独创性,也无法构成商业秘密,因为该数据条目可以从公开渠道获得。[40]国内也有类似意见认为:“对于现实中大多数的数据信息而言,信息制作者采集的信息本身大多来自公有领域,是任何人均可以从公开渠道直接获取的,显然,将各地为公众所知的信息汇编之后形成的成果认定为具有秘密性是荒谬的。”[41]这些意见实际上忽略了数据集合与它所包含的数据条目的秘密性的差别。数据集合整体的“普遍知悉”或“容易获得”,应该是指公众“普遍知悉”或“容易获得”数据集合的整体,而不是单个条目的“普遍知悉”或“容易获得”。单个数据条目处在公共领域,并不当然意味着数据集合整体就很“容易获得”。如果数据条目分散在公共领域,而公众将这些数据条目收集起来需要耗费实质的劳动和投入,则意味着该数据集合作为一个整体,不为公众所普遍知悉,也不容易获得,因此依然具有秘密性。当然,如果将公开的数据条目收集起来放在一起是很容易的事情,则收集者所得到的数据集合并不具备秘密性。

虽然笔者主张利用商业秘密法保护整体上依然

具有秘密性的数据集合，但并不反对通过专门立法保护整体公开的数据集合。如果收集者对外公开的数据条目达到实质数量并耗费收集者实质投入，则有可能、有必要给予有限的排他权(比如公开传播权)，以阻止部分不正当竞争行为。[42]专门立法对公开数据集合的有限保护，应该远比商业秘密法对未公开数据集合的保护力度要弱。它不是要取代商业秘密法，而是与之配套或衔接，弥补商业秘密法不保护公开数据的不足。

(二)无创造性但有实质投入的数据集合

在数据条目本身没有秘密性的情况下，如果收集者对数据条目进行了选择、编排和加工，体现了个人的创造性的劳动，则一般认为这些数据集合整体在公共领域并不存在("不普遍知悉")，且"不容易获得"，具备秘密性。如果收集者单纯付出了实质性的收集成本但未付出创造性劳动，则数据集合的秘密性可能存在争议。比如，网络平台将源自用户或物联网设备的数据条目汇总在一起形成的数据集合，类似中国期刊网或Google图书馆项目制作的数字化的作品数据集合等。部分意见倾向于在秘密性审查环节引入创造性的要求，即强调此类数据集合缺乏收集者的创造性劳动，虽然不被普遍知悉，但不满足"不容易获得"的要求。

表面上，商业秘密的构成要件中并没有直接的创造性的要求。但如果愿意，法院的确有可能通过解释秘密性要件中的"不容易获得"要求，使之涵盖创造性的要求。"不容易获得"，既可被解释为需要耗费实质性投入，所以不容易获得，也可被解释为需要付出创造性劳动，所以不容易获得。法院如果强调后者，则相当于在秘密性标准中变相地引入创造性的要求，对数据集合的秘密性认定有重大影响。

从商业秘密保护的司法实践看，各国法院在"创造性"问题上常常含糊其词，缺乏明确论述。美国法上，信息汇编类(compilations)商业秘密案件容易涉及这一问题。传统意义上的信息汇编，通常是指企业在经营过程中编制的体现客户偏好或个性需求的客户信息。[43]汇编结果体现了汇编人员直接的主观判断和选择。部分法院因此强调，数据收集者增加了他大脑里的东西(things from his head)，使得信息汇编具有了价值；[44]或者，商业秘密中融入了个人的分析判断。[45]但是，很少见到法院明确宣称，信息汇编人员的主观分析判断和选择编排是此类信息汇编具有秘密性的前提条件。

英国法院在信息汇编类商业秘密案件中也强调个人判断力，并接近引入创造性要求。在De Maudsley v. Palumbo案中，英国法院指出，将单个不具有新颖性的特征组合在一起，并不当然使得汇编结果具有新颖性(novel)。[46]这里所说的新颖性应该是指商业秘密法意义上的秘密性。英国著名法官拉迪(Laddie)认为，数据收集者如果只是付出机械劳动，可能并不足以使得数据集合具有秘密性(confidentiality)，该汇编行为应是人脑思维技巧的产物(the product of the skill of the human brain)。仅将公开可得的信息不加选择地编在一起，即使耗费一定的时间和精力，该信息集合也不应被视为具有秘密性(confidential)，因为没有应用相关的技巧(relevant skill)。[47]该法官认为，如果承认这类信息集合具有秘密性，就将导致他人可以自由获取信息条目，却不能将它们集中起来使用。[48]

在我国，最高人民法院并未在司法解释中明确"不容易获得"是否包含创造性的要求。《侵犯商业秘密规定》第四条列举了"有关信息为公众所知悉"的诸多示例，从该条的文字表述中看不出有创造性的要求。过去最高人民法院法官在解释秘密标准时，认为"那些相关人员不需要创造性劳动，仅仅是经过一定的联想即能获得的信息，就是容易获得的信息"。[49]这里似乎暗含某种创造性要求。地方法院在个别案例中对商业秘密提出创造性的要求，冯勇诉微软公司案就是一例。在该案中，冯勇通过比对微软拼音输入法中的汉字注音与公开出版物上的标准注音，发现了微软拼音输入法的诸多错误，并整理出校正清单。关于该校正清单是否属于商业秘密，法院认为："冯勇对拼音输入法中字的注音罗列后与公开出版物上字的注音进行对比后，从而发现其中部分注音不当，这种校正工作虽然工作量较大，但属简单的智力活动，不包含任何创造性智力劳动。不为

公众所知悉是指商业秘密应具有一定新颖性和创造性，即已经达到一定的技术水平，商业秘密与已有智力成果相比，必须具有一定的进步性，亦即该项技术秘密是创造性劳动的结果，而非本专业的一般技术人员不经研究就能够得出，也不是借助简单的推理和实验即可必然获得。”[50]

在数据集合的商业秘密保护成为热点问题之前，对于秘密性标准中是否应包含创造性要求，深入的学术研究并不多见。英国权威教科书的作者感叹，将公开信息转化为商业秘密法保护的信息集合，究竟需要何种类型和程度的劳动，是一个困难的问题，没有引起学术界足够的关注。[51]他们倾向于认为，数据集合的秘密性的门槛应该比较低，仅仅需要体现一定程度的技巧、劳动或判断(some degree of skill，labour or judgement)。[52]这一意见与英国法院的立场有明显的差异，似乎降低了对思维技巧或个人判断力的要求。在这一意义上，它优于上述拉迪法官的意见。

不过，笔者认为，我们应当比上述英国教科书的意见走得更远，彻底拥抱“实质性投入”标准，即商业秘密的秘密性标准中不应包含创造性的要求。只要数据收集者付出“实质性投入”，就满足了“不为公众所知”的秘密性标准。这里的“实质性投入”应当作相对宽泛的理解，既可能是量上的资本或劳动投入，也可能是质上的劳动投入，即体现了收集者的创造性劳动。

在认定商业秘密是否“容易获得”时，采用“实质性投入”标准符合商业秘密法设置秘密性要件的初衷。如前所述，商业秘密法设置秘密性要件的目的在于维护公共领域的行动自由，确保受保护的商业秘密能够为经营者带来竞争优势。为了实现上述立法目的，商业秘密法并没有必要刻意区分机械劳动(或单纯资本投入)和创造性劳动。只要数据集合需要耗费收集者实质性的机械劳动或单纯的资本投入，那就意味着公共领域并不存在现成的数据集合。对它进行保护，并不妨碍公众在公共领域的行动自由。同时，耗费了实质性投入通常意味着数据集合能够为经营者带来实质性的竞争优势。如果商业秘密法不保护这一类耗费实质性投入的劳动成果，就会使商业秘密法的立法目的受挫：纵容背信等违反商业道德的行为，增加企业间数据交易的成本，使得数据行业陷入丛林法则，刺激企业私下耗费更多资源于保密措施，造成社会资源的浪费。[53]最终，这会影响企业在数据服务领域的投资积极性。这原本正是商业秘密法所力图避免的结果。

实际上，传统商业秘密所保护的诸多秘密信息内容，很可能也只是持续努力经营过程中偶然原因成就，而未必一定体现最低限度的创造性。比如，客户信息、[54]合同报价、核心原材料来源信息等等就都是如此。在具体个案中，是否有创造性贡献也是法院很难事后查证的事实。因此，没有特别的理由要在数据集合的秘密性判断中引入创造性要求。在一些信息汇编类商业秘密案例中，中外部分法院强调信息汇编者的创造性劳动或个人判断力，在特定的时代背景下有一定的合理性。众所周知，传统的信息汇编类数据集合的规模比较小，数据的数量通常有限，仅仅依据数据规模这一事实常常不足以证明收集者付出了实质性的投入。比如，典型商业秘密案件中的客户信息的数据条目不过数条或数十条。[55]法院在判断秘密性时，强调汇编类商业秘密应体现个人的创造性劳动或判断力(质量)，[56]而不是强调数据收集的规模和汇总数据的行为本身的成本，是完全可以理解的。这大概也是在“以质取胜”的传统时代的合理选择。

不过，在面对更大规模的数据集合时，上述“以质取胜”的思路就显得不合时宜，应该代之以“质量并举”的思路。在二三十年前，移动存储设备的容量还相当有限，数据尚未像今天这样被大规模数字化。大规模的数据集合原本就不多，侵权也不容易。在依靠存储量只有2M的软盘交换数据，同时又没有有效分析工具的年代，大规模数据集合的商业秘密保护，并无现实需求。[57]因此，过去的商业秘密法没有经历太多的侵害大规模数据集合商业秘密的案件，[58]传统信息汇编类案件的判决思路还在顽强地发挥影响力。现在，网络数据行业发生了天翻地覆的变化。数据的大规模存储、自动整理加工、网络检

索等技术飞速发展。“海量数据的收集工作本身常常耗资巨大,同时,数据收集者不再需要以某种体现独创性的方式对数据进行深加工,就可以直接向用户提供数据。这是因为数据检索技术的进步使得很多信息的个性化整理、分类和编排变得不再重要。用户可以方便地在非结构化的数据集合中找到自己所需的具体信息。对于用户而言,重要的是数据本身,而非收集者本身的独创性贡献。”[59]在新的技术条件下,如果商业秘密法故步自封,继续强调收集者对数据选择和编排过程中的创造性劳动或个人判断力,就会导致比传统信息挖掘或汇编耗费更高成本也更有商业价值的数据集合反而无法得到商业秘密保护的尴尬局面。这明显违背了商业秘密法的立法目的。

美国有学者认为,在汇编作品上,著作权法放弃了“额头出汗”标准,商业秘密法应该与著作权法在这一问题上保持一致,也就是说,在考虑数据集合是否应该获得商业秘密保护时,数据收集过程中的实质投入是一个无关的因素。[60]这一意见忽略了著作权法与商业秘密法的重要差别。著作权法没有将“实质性投入”作为独创性的替代标准,是因为立法者认为在绝大多数情况下,著作权法保护作品中的独创性表达就可以为作者提供足够的激励,而无须进一步将保护延伸到不具有独创性的事实或抽象思想。从公众的角度看,这导致部分没有独创性但依然需要激励的“作品”(比如没有独创性的数据集合)无法获得著作权保护,需要在著作权法之外寻求补充性的保护。商业秘密保护正是在著作权法之外的补充保护选项。

将“实质性投入”标准应用到数据收集领域,绝大部分达到商业规模的数据集合,其收集过程都需要耗费实质性的成本,能够轻松满足这一标准。比如,网络平台利用平台系统自动收集的百万或千万用户的个人信息、交易数据、用户创作内容的集合,网络导航系统人工或自动收集的海量的地址信息、交通路线信息集合,网络搜索引擎利用网络爬虫收集的关于成千上万的网络站点内容的数据集合,轻易就能耗费收集者上百万元甚至更高的成本。因此,只要这类数据集合的文件包或其实质部分,并未被收集者完整对外提供,则此类数据集合应该能够轻松满足秘密性要求,即该数据集合整体上并“不容易获得”。当然,在具体个案中,我们依赖法院结合相关行业实践来确定诉争数据集合是否满足“不容易获得”标准所需要的实质性投入或数据规模。[61]有人可能会担心这一标准过于模糊,使得数据集合的客体边界不够明确。其实这是所有类型商业秘密保护都要面对的难题,并非数据集合所特有。想象一下,最为典型的技术信息和客户信息,关于它们是否“不容易获得”的判断,可能远比大规模数据集合要复杂。[62]我们并没有因此对商业秘密保护制度失去信心。

(三)数据条目受控开放的数据集合

笔者在前一节中关注的是,分散在公共领域或处在公开状态的数据条目,被集中起来形成数据集合后,该数据集合的秘密性。假定收集者在获得数据集合后,不再对外提供该数据集合的数据条目,因此,分析的重心就放在收集前数据条目的存在状态以及收集工作耗费的成本上,由此得出结论也很清楚,即这类数据集合通常能够满足秘密性的要求。

在本节中,笔者准备再往前迈一步,假定企业在收集完成数据集合后,因为自己商业模式的需要向公众公开提供数据条目。典型的做法是:企业将数据集合整体存储在网络服务器后台并采取访问口令或类似保密措施,未经特别许可,公众无法访问该后台数据集合;同时,收集者通过前台用户端界面,许可公众(用户)通过前台渠道获取一部分数据条目。比如,公交信息服务平台在对外提供公共汽车的行驶路线和时刻信息查询服务时,仅仅在前台向单一用户提供其查询的特定目标路线的相关公交信息,而不提供后台存储的实质量的公交数据信息。在提供此类信息时,平台会采用技术措施,限制公众绕过前台渠道直接访问后台存储的与本次查询服务无关的公交信息。[63]又如,著名的社交网站LinkedIn许可注册用户人工浏览其他用户公开发布的内容,但是在网站的用户协议中禁止用户使用自动工具从其服务器下载数据或规避访问限制措施。[64]同时,它也实

际采取技术措施阻止用户利用网络爬虫工具。[65]多数面向公众的数据服务大多采用类似模式，比如，地图导航服务、搜索引擎服务、期刊全文数据库服务、股市信息服务、天气预报服务，等等。这些服务商许可用户获得公开的碎片化的数量有限的数据条目信息，但是并不许可用户直接访问或下载后台存储的完整的数据集合。这一做法与上一节所假设的公开数据条目汇集成数据集合后被完全封闭起来的情形有重要的差别。完全封闭的数据集合容易满足秘密性要件，争议较小。企业通过前台限制性地提供数据条目，是否影响后台存储的数据集合的秘密性，则是一个充满争议的问题，值得深入探讨。

在探讨这一问题之前，有必要在此先回应部分人可能会有的另一疑惑：既然企业通过前台对外公开数据条目，那为什么还需要关注此类数据集合的秘密性问题呢？这是因为企业仅仅许可用户通过前台获取公开的数据条目，而不希望用户或竞争对手破坏它的后台的控制措施，直接下载后台存储的数据集合本身。可是，总是有部分用户或竞争对手试图突破这一限制，破解后台访问口令之类的加密措施，直接下载包含数据集合整体内容的文件包。这时，企业如果针对该非法获取后台数据集合的行为提起商业秘密侵权之诉，后台存储的数据集合是否具有秘密性，就成为这一争议中的关键问题。前述深圳公交汽车信息案就是这方面的典型案例。该案中，公众不能够获取原告存储在服务器后台关于公交实时情况的数据集合。被告采用黑客手段破解了原告客户端安装包的加密算法获得密钥，将自己的程序伪装成原告授权客户端，然后利用爬虫工具下载原告服务器后台数据，日均获得300万至400万条实时数据。[66]再如，在湖南蚁坊案中，被告通过技术手段破坏或者绕开新浪微博所作的技术限制，获取新浪服务器后台存储的用户登录后都难以看到的数据。[67]

在上述争议中，应当从观念上区分后台存储的数据集合和前台公开的数据条目。从商业秘密保护的角度看，二者可以是相互独立的客体。对于后台存储的数据集合，网络平台通常会采取有效的技术措施阻止公众未经许可访问。比如，设置API接口的访问密码；对数据包本身加密，避免泄露后被识别，等等。此类API接口或文件加密措施完全禁止用户未经许可以人工方式或通过自动工具访问数据集合。[68]未经特别授权，公众无法通过该接口获取该数据集合的任何内容；即便数据不慎泄露，公众通常也很难破解该加密措施而读出数据具体内容。因此，从商业秘密法的角度看，这类保密措施的合理性通常不会受到质疑。如果第三方违反许可协议，[69]或者破坏网络平台的后台保密措施，直接从后台下载存储了数据集合整体或实质部分的文件包，甚至破解平台针对该文件包采取的加密措施，则这一行为侵害了平台就该数据集合整体所享有的商业秘密。

对于网络平台通过前台客户端许可公众访问的数据条目，即便平台利用用户协议和反爬虫措施限制用户获得过多的数据条目，也不影响所有通过前台能访问的处在分散状态的数据条目本身被视为公开信息。理由是，网络平台虽然采取了反爬虫措施限制单个用户获得数据条目的数量，但是，只要任意一条数据条目处于可以通过人工浏览的方式获取的状态，则该数据条目本身就不再具有秘密性。用户规避反爬虫措施，或者违反禁止使用爬虫工具的约定，利用爬虫工具替代人工从前台渠道收集这些公开的数据条目，是否应当承担违约责任或其他法律责任，如侵权法、网络安全法、反不正当竞争法(包括专门的公开数据集合的保护立法)、反垄断法上的法律责任，已经超出本文关注的范围。即便用户依据这些法律需要承担法律责任，那也不是因为它侵害了商业秘密。以商业秘密保护的名义来制止用户使用爬虫工具获取原本公开的数据条目，不符合商业秘密法的内在逻辑。

接下来的问题是，平台通过前台向用户公开提供数据条目时，后台存储的数据集合整体的秘密性是否因此而受到影响呢？如前所述，这里假定网络平台虽然许可公众通过网页浏览或客户端界面获取数据集合中任意数据条目，但并不许可它们直接访问后台存储的数据集合整体(文件包)。从秘密性判断的角度看，后台存储的数据集合文件包是否具有

秘密性，不只取决于平台是否直接针对数据集合采取有效保密措施(比如访问口令限制与数据包加密)，还取决于公众通过前台爬取该数据条目并重新汇总的难度或成本大小，或者公众通过第三方公开渠道收集数据汇总后形成相同数据包的难度或成本大小。现实中，平台大多会采取直接的后台保密措施，同时也会在前台采用有效技术措施对抗网络爬虫。网络平台利用技术手段侦测到用户使用爬虫后，会立即停止向用户提供数据服务或采取其他限制措施。这导致公众通过前台爬取服务器数据的成本急剧上升，不再能轻易获得数据集合的实质内容或整体。公众要像收集者那样从其他公开渠道重新收集所有的数据条目，通常更加困难，耗费甚至超出原始的收集者。因此，虽然网络平台许可用户通过前台以受控制的方式获取有限的数据条目，可是用户要获取该数据集合整体或实质部分并不容易。此时，从法律上我们不应将前台界面上任意数据条目的公开性，解释为后来存储的数据集合整体的“普遍知悉”或“容易获得”，否则商业秘密法对于数据条目与数据集合秘密性的区分就会失去意义。

上述分析思路的假设前提是，网络平台采取有效的反爬虫措施，导致用户很难通过爬虫工具在极短的时间里以很低成本收集前台的数据条目，并将它们汇总起来形成新的数据集合。此时，反爬虫技术措施之于数据集合，类似于程序代码加密之于软件源代码。只要反爬虫措施达到合理的“强度”，能有效阻止绝大部分专业用户获得后台服务器中存储的实质数量的数据条目，就能够保证用户通过前台重新收集数据集合的内容变得“不容易”。换言之，前台零散的数据条目的公开，并不当然导致后台存储的数据集合本身失去秘密性。

如果平台采取有效反爬虫措施这一假设前提不存在，即用户很容易利用爬虫工具从前台获得与网络平台后台存储的数据集合内容大致相同的新的数据集合，则意味着后台存储的秘密信息实际上处在“容易获得”的状态，从而失去了秘密性。其中的道理就像企业采取保密措施保护自己存储在电脑里的产品设计方案，同时又公开出售自己的产品，公众很容易通过观察或测量该产品而获得该设计方案内容。这时候，即便第三方不当获取了企业存储在电脑里的设计方案，也未必侵害企业的商业秘密，原因是该设计方案事实上已经处在“容易获得”状态而被认为不具秘密性。

美国的CompuLife v. Newman案很好地说明了这一点。在该案中，CompuLife通过公开网络渠道收集了大量的保险公司的人身保险种类、服务条款和费率等信息，整理后存储在自己服务器上的数据库里，对外提供查询服务。保险公司更新数据后，CompuLife也会及时更新自己的数据库。用户通过网页递交查询关键词后，CompuLife通过网页反馈给用户查询结果，即用户所希望了解的保险服务的可能保险费率等信息。CompuLife利用技术措施将数据库保护起来，公众不能直接接触服务器上存储的该数据库集合本身。被告利用原告客户的账号访问其数据库，在很短时间里发出80万份“查询请求”(get commands)。每份“查询请求”模拟一个用户场景，原告服务器提供50家左右保险公司的服务条款。面对上述“查询请求”，原告的服务器提供了大约4350万份查询结果。这些结果信息被存入被告的服务器，然后对外提供跟原告相同的服务。原告在自己的信息中添加了水印，从而发现被告这一行为。原告得知上述行为后，并没有马上关闭服务器对“查询请求”的响应功能。原因是，其他程序员还是依赖这一简单的功能来获得信息。后来，原告添加了一项新的功能(degrade function)，可以有效阻止此类数据抓取行为(scraping)。当然，在权利人对公众开放查询服务的情况下，绝对禁止任何数据抓取行为并不现实。此外，原告事后还添加了用户协议，对数据抓取行为进行约束。不过，在本案被告借用用户账号抓取数据时，该用户其实并不受用户协议的约束。[70]

在该案中，美国第十一巡回上诉法院认为，尽管原告数据库中保险公司的单条报价信息是公开的，但是这并不意味着包含这些报价信息的数据库作为一个整体，不是商业秘密。被告从数据库中直接获取数据，到一定程度后就会侵害原告的商业秘密，否则法律明确规定的信息汇编(compilations)的商业秘

密保护就失去意义。在确认数据集合(数据库)构成商业秘密的基础上,法院认为原告预期自己用户通过人工浏览网页并检索的方式获得相关信息,而被告借用原告客户账户并利用黑客手段(爬虫技术)大量获取数据的方式,构成商业秘密法意义上的不当获取行为(misapproriation)。[71]

笔者认同美国法院关于数据条目来源于公开领域,并不妨碍该数据集合本身获得秘密性的论述;也支持法院关于原告许可用户通过网页检索界面获取任意数据条目并不当然妨碍该数据集合继续具备秘密性的暗示结论。但是,笔者并不认为原告对外提供数据服务时,有效控制了用户的浏览或检索行为。相反,用户在法律上并未受到保密协议约束(没有签署用户协议),在技术上也没有受到有效的反爬虫措施的制约,可以相对轻松地从原告的服务器上收集公开的数据条目,进而获取数据集合的实质部分内容。[72]从法院判决所披露的获取过程看,难谓"不容易获得"。因此,笔者倾向于认为,原告在提供数据服务时,该数据集合内容并不处于"不容易获得"的状态,因此不满足秘密性的要求。

四、商业秘密之外的替代选择

与笔者于本文中坚持尽量依靠商业秘密法保护企业数据集合的主张不同,很多意见主张在商业秘密法的框架外解决企业数据集合的产权保护问题。比如,日本和韩国的立法者就没有选择重新解释现有的商业秘密制度,使之能够覆盖大部分企业数据集合,而是选择在商业秘密之外,通过平行立法来保护部分企业数据集合。中国也有学者主张借鉴这一立法思路。[73]2022年11月国家市场监管总局公布的《中华人民共和国反不正当竞争法(修订草案征求意见稿)》第十八条关于商业数据的规定也明显受日韩立法思路的影响。此外,还有很多具备民法或网络法背景的学者在商业秘密法之外,支持统一的数据产权立法,赋予企业宽泛的数据权利。[74]接下来,笔者将逐一简要反驳这两类主张。

(一)变相的"限定提供数据"立法

依据日本《不正当竞争防止法》第二条第七款,如果权利人收集了合理数量的以电磁形式存储的技术或商业信息,仅仅向特定用户附条件持续提供,并施加了技术管理措施,则此类数据集合受到保护。这里数据集合被称作"限定提供数据"(shared data with limited access)。为了避免这一定义涵盖商业秘密,日本《不正当竞争防止法》第二条第一款第六项和第九项还明确规定,商业秘密被排除出"限定提供数据"的范围之外。依据日本的上述立法,他人不得通过盗窃、欺诈、胁迫或其他不当方式获取该数据集合或使用和披露前述不当获取的数据集合;也不得违反"管理条件"(the duties regarding the management)使用和披露受保护的数据集合。2021年,韩国《反不正当竞争和商业秘密保护法》第二条第一款第k项也引入类似的数据保护条款。

日本政府起草的《"限定提供数据"指南》(Guidelines on Shared Data With Limited Access)仔细说明了"限定提供数据"与商业秘密的关键差别是权利人是否有"保密"(keep confidential)意图。在该指南看来,企业和雇员签署协议,要求雇员不得对外提供秘密信息,这里有明显的"保密意图"。对于"限定提供数据",权利人采取电磁管理措施限制第三方接触,只是为了确保自己营利目的得以实现而非为了"保密",因为任何第三方只要符合权利人的营利目的就会被许可接触该数据。[75]该指南对于商业秘密的权利人的"保密"目的的理解,多少有些让人费解。商业秘密权利人向任何接触者或使用者发放许可,也都是因为该许可符合自身的营利目的。就许可目的而言,商业秘密许可与"限定提供数据"许可无法有效区分。从该指南关于商业秘密和非商业秘密信息的对比举例看,[76]日本决策者似乎相信,商业秘密权利人采取保密措施后,只能非常谨慎地授权为数不多的用户接触该商业秘密。这大概是该指南认为雇员保密协议体现雇主"保密"意图的原因。如果获得授权持续接触数据的人员过多,则不再有所谓的"保密意图",只能视为"限定提供数据"而不是"商业秘密"了。其实,仅仅依据接触者人数的多寡来区分商业秘密与非商业秘密("限定提供数据"),并非合理的选择。

近一百年前,美国最高法院在Board of Trade案

中就考虑过类似的商业秘密许可人数过多的问题。在该案中,权利人收集了谷物交易市场上的实时价格信息,然后通过电报传送到全国各地众多的客户办公室。法院认为,即便很多人和收集者签署合同而接触了此类秘密信息,只要每一个接触者都签署了保密协议,就不妨碍法院认定其为商业秘密。[77]不过,笔者认为,这里法院还是要考虑数据的数量和性质,以及公众接触的程度。在某些情况下,数据收集者向太多人提供数据,即便所有接触者都签署保密协议,收集者也很难有效监督他们以保证保密协议得到有效执行,法院亦有可能认定该数据事实上丧失秘密性。[78]在商业秘密是相对简短的信息时尤其如此。比如,对于产品的配方、加工工艺中的关键参数等,如果权利人授权业内成千上万的自然人接触该配方或了解该关键参数,则即便这些人都声明承担保密义务,也很难在事实上阻止该工艺参数在相关行业变成众所周知的事实。在众多被许可人事实上随意传播相关信息的情况下,商业秘密法许可权利人事后选择性地追究责任,会使得被告在竞争中处于不利地位——相关信息事实上已经公开,自己却要将它视为商业秘密。不只如此,这还会导致很多人无意中陷入各种纠纷链条,面对不确定的法律风险。这实际上损害了公众在公共领域的行动自由。因此,这时候,法院认定此类接触者众多的信息不再具有秘密性,或者认定单纯的保密协议并非合理的保密措施,是可以理解的。

不过,大规模的数据集合与产品配方或制作工艺方面的秘密信息有巨大区别,数据集合的数据条目的接触者增多,并不当然导致公众更容易获得该数据集合的实质部分。如果平台采用技术措施,仅许可用户从前台获取数据集合的有限条目,禁止从后台获得数据集合的实质内容,则即便通过前台接触数据的用户众多,也不妨碍该后台存储的数据集合本身的秘密性。比如,在淘宝诉美景案中,淘宝的"生意参谋"数据集合包含淘宝平台海量商品的销售情况和店铺的经营情况,数据量巨大。普通用户通过许可的用户界面检索和浏览,并不能使用爬虫工具,根本不可能获得数据集合的实质内容。同时,淘宝要求每个用户都承担保密义务。因此,即便该产品的授权用户超过2000万,月服务商家超过500万,公众依然无法轻易下载数据集合的整体或实质内容。[79]因此,该数据集合整体依然处在"难以获得"的状态,即满足所谓"秘密性"要求。当然,如果淘宝授权每个用户获得完整的数据库拷贝,则即便每一个用户都签署了保密协议,此等规模的许可也的确有可能导致该数据库丧失商业秘密法上的"秘密性"。由此看来,判断数据集合是否具有秘密性,关键不在于接触者数量的多少,而在于特定的接触规模是否事实上导致公众很容易获得该数据集合的实质部分。如果技术措施有效阻止第三方公众获取数据集合的实质内容,接触人数多少就不再是问题。另外,按照接触者的数量来区分商业秘密与"限定提供数据",还会产生中间界限如何划分的难题。

当然,从日本法"限定提供数据"定义对于数据规模("合理数量")的要求,以及《"限定提供数据"指南》推定它为大数据或类似数据的表述看,[80]日本决策者显然认为数据规模大,也是"限定提供数据"与传统商业秘密的一大区别。在笔者看来,数据规模大并非"限定提供数据"与商业秘密的本质区别。相反,数据规模大到一定程度后,反而有利于证明该数据集合耗费实质收集成本,不容易从公共领域获得,因而整体上更有可能具有"秘密性"。

既然日本法上所谓"限定提供数据"与商业秘密并无本质区别或截然界限,立法者在商业秘密法之外,为"限定提供数据"制定平行的保护规则,则必然会出现浪费资源重复立法的问题。日本《不正当竞争防止法》关于"限定提供数据"的客体定义、权利内容条文,与商业秘密保护条文其实惊人的相似,明显昭示着立法者在叠床架屋。[81]这里可以从操作层面日本法对所谓"技术管理措施"的要求进一步说明这一问题。依据日本《"限定提供数据"指南》,技术管理措施要能够阻止任何未经许可的第三方接触该数据,包括账户ID和密码控制、限定访问终端、生物身份认证措施(比如脸部识别与指纹认证)、数据加密等,旨在让第三方了解权利人的保护意图和受保护数据的范围。[82]这里所列举的技术管理措施,与商业

秘密法下的保密措施并无本质差别。这些措施使得未经授权的第三方不能获取“限定提供数据”的集合或局部条目。可以想见,相当一部分“限定提供数据”事实上处在公众难以获取的状态,整体上将它视为商业秘密保护的客体并无障碍。第三方破坏访问口令等技术管理措施,直接获取服务器上存储数据集合的文件或数据包,按照商业秘密侵权来处理也很顺畅。因此,在商业秘密法之外,为整体上并未对外公开的数据集合创设新的特殊保护规则,并无必要。

平行立法的危害并不只限于浪费立法资源,更糟糕的是,它会在两部法律衔接地带人为制造模糊性,导致具体个案中出现无谓的法律适用的争议。如前所述,在日本法下,“限定提供数据”与商业秘密之间的界限并不明确。比如,小规模的客户信息(客户名单)、合同报价信息被认为是传统的商业秘密;而大规模的客户交易信息集合或用户个人信息集合就可能被认为是“限定提供数据”。理论上很难说清楚二者之间的界限究竟在哪里。如果考虑到二者的保护力度实际上很接近,就更难理解为什么要在二者之间武断地划线。从实用主义的角度看,决策者还不如直接承认,“限定提供数据”中大部分实际上就是商业秘密,合理解释或完善现有商业秘密法就能解决问题,何苦要重新立法去建立一套不成熟的平行体系呢?当然,对于不符合商业秘密标准的“限定提供数据”即公开的数据集合,有可能需要通过专门立法提供有限的保护。遗憾的是,对于此类数据集合,日本的“限定提供数据”规则又明显给予过于宽泛的保护,同时提出技术管理措施要求,同样不是很好的立法例。[83]在笔者看来,日本这一立法处在两头不讨好的尴尬境地。

(二)无秘密性要求的“统一数据产权”

《中华人民共和国民法典》和《数据二十条》笼统地使用“数据”概念,并未区分秘密与公开数据。很多建议统一数据产权立法的学术意见,也没有充分考虑现有的二分思路,而是笼统地建议赋予企业对数据的占有(持有)、使用、处分(处置)和收益的权利(或其他类似的权能)。[84]这给人的印象是,未来的统一数据产权立法可以不区分秘密数据与公开数据。其实,公开数据和秘密数据的保护需求和保护模式存在巨大差异,不加区分地统一立法是不现实的。

对于公开数据集合,企业通常并不需要宽泛的产权保护。在多数情况下,企业只有在自己的商业模式能够确保它获得合理的回报时,才会公开数据。对它们而言,法律对公开数据集合的宽泛产权保护是多余的。只有在少数情况下,如数据收集工作耗费企业实质投资,而竞争对手复制并直接与自己竞争时,才可能需要法律提供有限的保护,比如禁止竞争对手以相同方式公开传播该数据(即赋予有限的公开传播权)。除此之外,他人通过公开途径获取数据以及后续私下利用或非竞争性的二次利用该数据的行为,通常不需要被禁止。比如,在美国有名的LinkedIn案中,法院就明显不认为下载社交媒体数据进行二次利用会直接损害社交平台的利益。[85]再如,企业通过网络爬虫从公开途径收集大量数据后训练自己的AI系统,只要不直接公开提供其收集的训练数据,通常也不会威胁数据源头企业的利益。从公众的角度看,如果法律对公开数据集合给予宽泛的产权保护,还会损害公众利用公共领域数据的自由,增加后续创新成本。这正是著作权法拒绝保护事实或技术类数据的原因所在。数据产权立法背离著作权法的基本原则,保护企业的公开数据集合,则必须对保护客体提出苛刻的入门要件,同时赋予很窄的排他性权利,从而将限制公众自由的负面影响压缩到可以接受的范围内。[86]

与公开数据相对的是企业采取保密措施的秘密数据。显然,企业希望法律能够帮助维持此类数据的秘密状态,以获得合理回报。按照商业秘密保护的正当性理论,如果法律不禁止公众破坏企业保密的努力,就会导致社会资源的浪费或企业投资积极性受挫:企业要么增加投资以提升自己保密措施的强度;要么因为无法获得有效保护,而降低收集数据的投资力度。正因为如此,《反不正当竞争法》第九条禁止任何人违反保密义务或破坏保密措施,不正当获取、披露、使用或允许他人使用该商业秘密。商业秘密法对于秘密数据的保护比未来法律对公开数

据的保护要宽泛得多，但并不直接威胁到公共领域的自由。这是因为商业秘密法上的秘密性和保密措施要求，确保它仅仅保护未进入公共领域的数据。因此，虽然商业秘密法所赋予的权能相对宽泛，但公众依然可以容忍。

既然秘密的和公开的数据集合有不同的保护需求，立法者要维持精细的利益平衡关系，就必然分别设计出不同的保护模式，使得两者在保护客体要件、权利内容、权利限制等重要方面出现明显差异。在这一背景下，如果立法者依然要统一数据产权立法，则只能选择更偏向秘密数据或公开数据持有者的预期，这样必然会顾此失彼(理论上还存在中间道路，将保护强度设置在商业秘密保护与公开数据保护之间，不过这样会更糟，两边不讨好)。

如果统一立法选择向商业秘密的保护水平看齐，则意味着统一立法放弃保密性或保密措施要求，不论数据公开与否，都赋予“禁止复制、使用或传播”之类的权利，或者类似民法学者所说的“占有(持有)、使用、处分(处置)和收益”之类的宽泛权能。[17]放弃商业秘密要件后，统一立法大概率需要利用“实质投资”或“实质数量”等模糊标准来界定受保护的客体，[18]界定权利边界的成本急剧增加。同时，失去合理保密措施这道门槛，公开数据的宽泛的权能将大大压缩公共领域，损害后续创新。到目前为止，宽泛的统一数据产权立法依然是不能想象的。

如果统一立法选择向公开数据的保护水平看齐，仅仅在数据集合满足实质投资和实质数量要件时才赋予有限的排他权，比如前述公开传播权，则统一立法无法满足那些原本选择对数据保密的企业的预期。失望的企业肯定会在统一立法之外，继续谋求商业秘密保护。如果统一立法之外的商业秘密保护主张得不到支持，企业会强化自助措施，这将导致商业秘密法原本要避免的社会资源浪费或投资激励不足问题“重现江湖”。如果它们的主张得到支持，则会出现低保护水平的统一数据产权立法与商业秘密法并存的局面。这实际上意味着，统一数据产权立法与商业秘密法重叠的部分，不过是一纸空文。这样的数据产权立法，实际上并未实现统一，而是变相地回归商业秘密法保护秘密数据而特殊立法保护公开数据集合的二分路径。

五、结论

在大数据和人工智能大行其道的网络时代，企业数据产权保护成为社会关注的焦点。对于企业数据，决策者应坚持公开数据与秘密数据二分的成熟思路，分门别类地进行保护。将企业秘密数据集合纳入现有商业秘密法的框架，即“新酒入旧瓶”，就能解决大部分数据产权问题。网络平台在生产经营过程中收集的数据集合，无论其所含的数据条目是个人信息、文学艺术作品还是用户交易数据，均可视为商业秘密法上的“经营信息”。数据集合中的数据条目来源于公共领域，也不妨碍该数据集合整体上获得秘密性。网络平台对后台存储的数据集合采用访问口令和类似有效的加密措施后，只要公众难以通过网络爬虫等工具从整体上获取和利用该数据集合的实质部分，则该后台存储的数据集合依然满足商业秘密法对秘密性的要求。

当然，现有商业秘密法上的权利归属、转让、许可使用、权利限制、侵权认定等重要规则尚不够完善。利用现有商业秘密法来保护企业数据集合，决策者还需要从上述多个方面作适应性的解释或改革。尽管如此，在商业秘密保护法之外，类似日韩那样进行所谓“限定提供数据”的平行立法，不过是叠床架屋，没有必要；忽视公开数据与秘密数据的重要区别，进行统一数据产权立法的思路，更不可行，不值得认真对待。总之，决策者只有深入理解商业秘密法的立法逻辑和政策目标，理解它在保护企业数据方面的重要角色，才能避免浪费立法资源去制造新的法律冲突。对于企业数据保护，更合理的选择依然是回到现有的商业秘密保护法加可能的公开数据特殊保护立法的思路。

注释：

①本文中的“商业秘密保护法”或“商业秘密法”并非特指现有专门立法，而是指商业秘密保护的相关法律制度。在现阶段的中国其主要指《反不正当竞争法》中商业秘密保护规则，将来则应该是指商业秘密保护方面的专门立法。

②有学者持相反意见，认为“商业秘密制度只能对大数据中的秘密数据提供‘有限保护’，且是以私力保密措施为基础的‘防御性保护’，难以承载数据产权保护的制度功能”。吴汉东：《数据财产赋权的立法选择》，载《法律科学》2023年第4期。这一意见忽略了绝大多数企业数据其实处在秘密状态的事实，也低估了企业对于“私力保密措施”的信任和依赖。多数私下的数据使用行为很难追踪，多数企业不会因信任法律保护而公开自己的核心数据。

③依据《反不正当竞争法》第九条第一款，数据资源生产者(持有人)可以禁止任何人以盗窃、电子侵入等不正当手段获取该数据内容，禁止任何人披露、使用或者允许他人使用以前项手段获取的数据内容，也可以禁止任何人违反保密义务或违反权利人保密要求，披露、使用或允许他人使用其掌握的数据内容。学术界对于商业秘密法所保护的究竟是一种财产性权利，还是合同权益，抑或是反不正当竞争法上的利益，尚存在学术争议。参见 Mark Lemley, The Surprising Virtues of Treating Trade Secret Rights as IP Rights, 61 Stanford Law Review 311, 319-329(2008)；唐海滨、孙才森、梁彦、王莉萍：《有关商业秘密立法的重点难点问题》，载《中国法学》1999年第4期。限于本文写作目的，笔者无意卷入这一差不多已绵延一个世纪的理论争议，而是使用“商业秘密权”或“商业秘密所有权”的表述单纯指代商业秘密法所保护的原始数据的收集者所享有的权益。

④在许可的范围内，数据加工者可以自由使用、许可第三方使用或对外披露自己的“数据衍生产品”。如果加工者实质改变数据内容后，通常也能对数据衍生产品主张新的知识产权保护。

⑤这一客体审查要件常常被人们忽略。比如，纯粹的文学艺术作品就很可能就因为不符合这一暗含条件而被排除出商业秘密的范围。关于这一客体审查要件的重要性，可以参见 Eric E. Johnson, Trade Secret Subject Matter, 33 Hamline Law Review 545, 577(2010)。

⑥参见梅夏英：《企业数据权益原论：从财产到控制》，载《中外法学》2021年第5期。

⑦See Peter A. Luccarelli Jr., The Supremacy of Federal Copyright Law over State Trade Secret Law for Copyrightable Computer Programs Marked with a Copyright Notice, 3 Computer Law Journal 19(1981-1982).

⑧See Charles Tait Graves, Caifornia´s Film Script Cases & Trade Secret Law, 44 Columbia Journal of Law & Arts 21, 64 (2020). 该学者仔细研究电影行业的剧本交易后认为，没有理由排除剧本、故事情节等要素的商业秘密属性。

⑨Robert G. Bone, A New Look at Trade Secret Law: Doctrine in Search of Justification, 86 California Law Review 241, 248 (1998).

⑩Amy Kapczynski, The Public History of Trade Secrets, 55 UC Davis Law Review 1367, 1391(2022).

⑪有相反意见强调商业秘密条款中的示例实际上起到限制商业秘密范围的作用，因而商业秘密客体的范围可能比想象的要窄很多。See Eric E. Johnson, Trade Secret Subject Matter, 33 Hamline Law Review 563(2010). 不过，这似乎并未成为美国学者的主流意见。

⑫See 18 U. S. C. § 1836(The Defend Trade Secrets Act)。该条还列举了一些具体的例子，包括模式(pattern)、计划(plans)、信息汇编(compilations)、程序装置(program devices)、配方(formulas)、原型(prototypes)、方法、技巧(techniques)、过程(processes)、流程(procedures)、程序(programs)或代码(codes)。

⑬Amy Kapczynski, The Public History of Trade Secrets, 55 UC Davis Law Review 408(2022).

⑭林广海等：《〈最高人民法院关于审理侵犯商业秘密民事案件适用法律若干问题的规定〉的理解与适用》，载《法律适用》2021年第4期。

⑮正如2020年《最高人民法院关于审理侵犯商业秘密民事案件适用法律若干问题的规定》第一条第三款所说，“客户信息，包括客户的名称、地址、联系方式以及交易习惯、意向、内容等信息”。

⑯参见北京微梦创科网络技术有限公司诉北京淘友天下技术有限公司等不正当竞争纠纷案，北京知识产权法院(2016)京73民终588号民事判决书。

⑰Marc A. Rodwin, Patient Data: Property, Privacy & the Public Interest, 36 American Journal of Law & Medicine 586, 588 (2010).

⑱参见深圳市谷米科技有限公司诉武汉元光科技有限公司等不正当竞争纠纷案，广东省深圳市中级人民法院(2017)粤03民初822号民事判决书。

⑲参见上海汉涛信息咨询有限公司诉北京百度网讯科技有限公司等不正当竞争纠纷案，上海知识产权法院(2016)沪73民终第242号民事判决书。

⑳商业秘密法限制对商业秘密的“使用”，包括功能性使用或私人使用。这可能超出了著作权法的保护范围。不过，对于权利人而言，这类使用的负面影响很小。因此，商业秘密法的补充保护没有特别的意义。

㉑参见英国案例 De Maudsley v. Palumbo[1996]FSR 447。该案中，关于新舞厅的布局和经营思路被认为过于模糊，无法获得商业秘密保护。

㉒17 U. S. C. § 301(a)。

㉓Spear Mktg., Inc. v. Bancorpsouth Bank, 844 F. 3d 464 (5th Cir. 2016); Globe Ranger Corp. v. Software AG U. S., Inc., 836 F. 3d 477(5th Cir. 2016).

㉔Kewanee Oil Co. v. Bicron Corp., 416 U. S. 470,491 (1974).

㉕Computer Assoc. Int´l v. Altai, Inc., 982 F. 2d 693, 716 (2d Cir. 1992).

㉖John M. Williamson, The Defend Trade Secrets Act and Copyright Preemption, https://www.finnegan.com/en/insights/articles/the-defendtrade-secrets-act-and-copyright-preemption.html,2023年8月8日访问。

㉗关于生产数据条目的成本与数据集合的收集成本的区分及其法律意义的深入探讨,参见崔国斌:《公开数据集合法律保护的客体要件》,载《知识产权》2022年第4期。

㉘比如,在衢州万联网络技术有限公司与周慧民等侵害商业秘密纠纷上诉案中,原告的网站经过三年的经营,在2006年时具有55万注册用户。法院认为,“55万注册用户的用户信息(包括用户名字段、注册密码字段和注册时间字段等信息)是无法从公开的渠道或采取简单的编排手段轻易获取的”,因此构成商业秘密。上海高级人民法院(2011)沪高民三(知)终字第100号民事判决书。

㉙关于商业秘密保护的目的,可以参考David D. Friedman, William M. Landes & Richard A. Ponser, Some Economics of Trade Secret Law, 5 Journal of Economic Perspectives 61 (1991); Mark Lemley, The Surprising Virtues of Treating Trade Secret Rights as IP Rights, 61 Stanford Law Review 311, 319-329 (2008); Robert G. Bone, A New Look at Trade Secret Law: Doctrine in Search of Justification, 86 California Law Review 241, 248 (1998)。

㉚比如,有学者认为匿名化的个人信息集合可能构成商业秘密,但是非匿名化的个人信息集合很可能不是收集者的商业秘密。Gintarè Surblytè, Data Mobility at the Intersection of Data, Trade Secret Protection and the Mobility of Employees in the Digital Economy, Max Planck Institute for Innovation and Competition Research Paper No. 16-03(2016), p. 18, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2752989,2023年8月28日访问。国内也有学者认为,源于个人的数据难以构成商业秘密。参见申卫星:《论数据用益权》,载《中国社会科学》2020年第11期。

㉛WTO Report of the Panel DS 290: European Communities-Protection of Trademarks and Geographical Indications for Agricultural Products and Foodstuffs, WT/DS290/R, p. 67, para. 7. 246.该裁决明确指出《与贸易有关的知识产权协议》(TRIPS协议)所要求保护的知识产权只是消极的排他权利,而非权利人自己积极实施的权利。

㉜蒋志培、孔祥俊、王永昌:《〈关于审理不正当竞争民事案件应用法律若干问题的解释〉的理解与适用》,载《法律适用》2007年第3期。

㉝Robert G. Bone, A New Look at Trade Secret Law: Doctrine in Search of Justification, 86 California Law Review 241, 249 (1998).

㉞崔国斌:《大数据有限排他权的基础理论》,载《法学研究》2019年第5期。

㉟Josef Drexl, Designing Competitive Markets for Industrial Data Between Propertisation and Access, Max Planck Institute for Innovation and Competition Research Paper No. 16-13(2016), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2862975, p.23, last visited on Aug. 28, 2023.

㊱参见申卫星:《论数据用益权》,载《中国社会科学》2020年第11期。与德国的德莱克瑟教授一样,申卫星教授似乎因为场所本身的公开性而否认特定收集者本身所得信息的秘密性。这并非商业秘密法判断秘密性的思路。耗费实质资源对公开售卖的商品进行反向工程所获得的技术信息,依然有可能成为该收集者的商业秘密。

㊲Tommaso Fia, Resisting IP Overexpansion: The Case of Trade Secret Protection of Non-Personal Data, 53 International Review of Intellectual Property and Competition Law 917, 926 (2022).

㊳Marc A. Rodwin, Patient Data: Propert, Privacy & the Public Interest, 36 American Journal of Law & Medicine 586, 588 (2010).

㊴类似结论参见Tommaso Fia, Resisting IP Overexpansion: The Case of Trade Secret Protection of Non-Personal Data, 53 International Review of Intellectual Property and Competition Law 917, 927(2022)。不过,作者认为这是知识产权过度扩张的表现,建议依靠知识产权法定原则等将这些未经加工的原始数据集合排除出商业秘密的保护范围。

㊵Sharon K. Sandeen, A Contract by Any Other Name is Still a Contract: Examining the Effectiveness of Trade Secret Clauses to Protect Databases, 45 IDEA-the Law Review of the Franklin Pierce Center for Intellectual Property 119, 134(2005).

㊶芮文彪、李国泉、杨馥宇:《数据信息的知识产权保护模式探析》,载《电子知识产权》2015年第4期。

㊷进一步的讨论超出本文的范围,可参考崔国斌:《大数据有限排他权的基础理论》,载《法学研究》2019年第5期。

㊸Sea Coast Fire, Inc. v. Triangle Fire, Inc., 170 So. 3d 804

(Fla. 3d DCA 2014); PSC, S. A. v. PriceSmart, Inc., 07-21383-CIV, 2007 WL2781021(S. D. Fla. Sept. 19, 2007).

㊹ Airfacts, Inc. v. De Amezaga, 909 F. 3d 84,96(4th Cir. 2018).

㊺Motor City Bagels, L. L. C. v. Am. Bagel Co., 50 F. Supp. 2d 460, 473-79(D. Md. 1999).

㊻De Maudsley v. Palumbo[1996]FSR 447, 459.

㊼英国的拉迪(Laddie)法官持此类意见。Ocular Science v. Aspect Vision, [1997]RPC 289, 374-375.

㊽ Ocular Science v Aspect Vision, [1997]RPC 289, 374-375. 这一担心表明,该法院似乎误解了商业秘密保护的本质。实际上,商业秘密法保护数据集合整体,只是禁止公众违反保密义务或者通过不当手段利用从权利人那里获得的数据集合的整体或实质部分,并不妨碍公众独立收集和利用具有相同内容的数据条目和数据集合。在权利人对数据集合整体采取保密措施但通过前台以受控方式对外公开提供有限数据条目的情况下,公众从前台合法获得数据条目形成数据集合后,利用该数据集合,一般也不受商业秘密保护的影响。进一步讨论可以参考本文第三部分第三小节。

㊾蒋志培、孔祥俊、王永昌:《〈关于审理不正当竞争民事案件应用法律若干问题的解释〉的理解与适用》,载《法律适用》2007年第3期。

㊿冯勇诉微软(中国)有限公司案,湖北省武汉市中级人民院(2003)武知初字第70号民事判决书。

(51) Lionel Bently, Brad Sherman, Dev Gangjee & Phillip Johnson, Intellectual Property Law, Oxford University Press, 2018, p. 1232.

(52) Lionel Bently, Brad Sherman, Dev Gangjee & Phillip Johnson, Intellectual Property Law, Oxford University Press, 2018, p. 1232.

(53) Mark Lemley, The Surprising Virtues of Treating Trade Secret Rights as IP Rights, 61 Stanford Law Review 311, 333-334 (2008).

(54)法院在判断客户信息的秘密性时,关注的核心要素就是原告"为客户信息形成所付出的劳动、金钱和努力"。参见江苏省高级人民法院《侵犯商业秘密民事纠纷案件审理指南》(2021年)第2.5.3节。

(55)在Art & Cook,Inc. v. Haber案中,诉争的客户信息只有70条。Art & Cook, Inc. v. Haberm, 416 F. Supp. 3d 191(E. D. N. Y. 2017).

(56)Sea Coast Fire, Inc. v. Triangle Fire, Inc., 170 So. 3d 804, 808(Fla. 3d DCA 2014); Kavanaugh v. Stump, 592 So. 2d 1231, 1232(Fla. 5th DCA 1992); E. Colonial Refuse Serv., Inc. v. Velocci, 416 So. 2d 1276, 1278(Fla. 5th DCA 1982).

(57)有学者将网络或电子数据集合类比为"现实生活中广泛存在的排斥他人访问的藏书馆、资料库等",认为"相关机构对其所控制的整体信息从未产生类似商业秘密保护的问题"。因此,企业数据集合并不适宜定性为商业秘密。参见梅夏英:《企业数据权益原论:从财产到控制》,载《中外法学》2021年第5期。其实,对于非电子化的资料库或藏书馆,未经许可侵入馆内不当获取实质数量的图书资料的情形,几乎不可能发生,即便发生也需要耗费入侵者实质性的复制成本,并且,收集者基于物理场所或图书财产的实际控制,也能获得救济。这应该是这一领域没有产生商业秘密保护问题的原因。如果技术进步到入侵者一夜之间可以合法地以极低成本物理复制一座藏书馆或资料库,然后与该藏书馆和资料库竞争客户,法律的应对就可想而知了,而这正是今天网络环境下数据收集者所面对的挑战。

(58)到目前为止,在美国和欧洲,真正处理大规模数据集合商业秘密争议的案件并不多见。参见Tommaso Fia, Resisting IP Overexpansion: The Case of Trade Secret Protection of Non-Personal Data, 53 International Review of Intellectual Property and Competition Law 917, 923(2022)。中国虽然涉及数据集合的争议较多,但这些争议也极少被视为商业秘密争议。

(59)崔国斌:《大数据有限排他权的基础理论》,载《法学研究》2019年第5期。

(60) Sharon K. Sandeen, A Contract by Any Other Name is Still a Contract: Examining the Effectiveness of Trade Secret Clauses to Protect Databases, 45 the Law Review of the Franklin Pierce Center for Intellectual Property 119, 137(2005).

(61)具体可以参考公开数据集合保护所需要的实质投入标准并作适当变通。限于篇幅,笔者于本文中不作深入讨论。参见崔国斌:《公开数据集合法律保护的客体要件》,载《知识产权》2022年第4期。

(62)以客户信息为例,法院判断客户信息的秘密性时要考虑诸多不确定的因素,包括原告"为客户信息形成所付出的劳动、金钱和努力"。参见江苏省高级人民法院《侵犯商业秘密民事纠纷案件审理指南》(2021年)第2.5.3节。

(63)参见深圳市谷米科技有限公司诉武汉元光科技有限公司等不正当竞争纠纷案,广东省深圳市中级人民法院(2017)粤03民初822号民事判决书。

(64) Article 8.2 of the LinkedIn User Agreement of 2022, https://www.linkedin.com/legal/user-agreement, 2023年8月28日访问。

(65) LinkedIn, Prohibited Software and Extensions, https://www.linkedin.com/help/linkedin/answer/a1341387/prohibited-

software- and- extensions?src=re- other&veh=www.natlawreview.com，2023年8月28日访问。

㊻参见深圳市谷米科技有限公司诉武汉元光科技有限公司等不正当竞争纠纷案，广东省深圳市中级人民法院(2017)粤03民初822号民事判决书。

㊼参见湖南蚁坊软件股份有限公司与北京微梦创科网络技术有限公司不正当竞争纠纷案，北京市高级人民法院(2019)京73民终3789号民事判决书。“微梦公司认为直接攻击微博平台服务器是蚁坊公司抓取微博平台后端数据最为可能采用的方式。”蚁坊公司在二审中试图提交证据证明，自己为政府提供舆情监测服务，取得了新浪微博运营方的同意通过五个数据接口账号获取和使用新浪微博的高权限数据。但该案中法院并未最终查明这一事实。即便后一说法属实，这也是违反了协议目的保留并使用微博后台数据。

㊽比如，Google公司就建议，如果不希望爬虫访问隐私文件，应采用文件加密方式加以控制。Google搜索中心：《robots.txt简介》，https://developers.google.com/search/docs/crawling-indexing/robots/intro?hl=zh-cn，2022年12月1日访问。

㊾比如，北京微梦创科网络技术有限公司诉北京淘友天下技术有限公司等不正当竞争纠纷案，北京知识产权法院(2016)京73民终588号民事判决书。在本案中，被告获得许可利用不对外开放的API接口获取微博的秘密的用户数据集合，但被告超出许可范围爬取更多的数据。这应该可以按照商业秘密许可争议来处理，但是法院并没有这么做。

㊿CompuLife Software, Inc. v. Newman, Case No. 9: 16-CV-81942-Rosenberg/Brannon, at 3-8(S. D Fla. Jun. 12, 2017).

(71)Compulife Software Inc. v. Newman, 959 F. 3d 1288, 1314 (11th Cir. 2020).

(72)CompuLife Software, Inc. v. Newman, Case No. 9: 16-CV-81942-Rosenberg/Brannon, at 3-8(S. D Fla. Jun. 12, 2017).

(73)孔祥俊：《论反不正当竞争法“商业数据专条”的建构——落实中央关于数据产权制度顶层设计的一种方案》，载《东方法学》2022年第5期。

(74)参见申卫星：《论数据用益权》，载《中国社会科学》2020年第11期；王利明：《数据何以确权》，载《法学研究》2023年第4期；张新宝：《产权结构性分置下的数据权利配置》，载《环球法律评论》2023年第4期。

(75)Japanese Ministry of Economy, Trade and Industry, Guidelines on Shared Data With Limited Access, Jan. 21, 2019, p. 11.

(76)Japanese Ministry of Economy, Trade and Industry, Guidelines on Shared Data With Limited Access, Jan. 21, 2019, p. 12.

(77)Board of Trade v. Christie Grain & Stock Co., 198 U. S. 236, 250-251(1905).

(78)Kristen Osenga, Information May Want to Be Free, But Information Products Do Not: Protecting and Facilitating Transactions in Information Products, 30 Cardozo Law Reviwe 2099, 2117 (2009); Jacqueline Lipton, Balancing Private Rights and Public Policies: Reconceptualizing Property in Databases, 18 Berkeley Technology Law Journal 773, 818(2003).

(79)淘宝(中国)软件有限公司诉安徽美景信息科技有限公司不正当纠纷案，浙江省杭州市中级人民法院(2018)浙01民终7312号民事判决书。

(80)Japanese Ministry of Economy, Trade and Industry, Guidelines on Shared Data With Limited Access, Jan. 21, 2019, p. 6.

(81)如前所述，2022年公布的《中华人民共和国反不正当竞争法(修订草案征求意见稿)》第十八条关于商业数据保护的内容，也受日本立法思路的影响。对照它与该法的商业秘密保护条款，我们也会发现两者惊人的相似。

(82)Japanese Ministry of Economy, Trade and Industry, Guidelines on Shared Data With Limited Access, Jan. 21, 2019, p. 8-9.

(83)从公开数据集合保护角度对日本立法的批评，参见崔国斌：《公开数据集合法律保护的客体要件》，载《知识产权》2022年第4期。

(84)龙卫球：《再论企业数据保护的财产权化路径》，载《东方法学》2018年第3期；许可：《数据保护的三重进路——评新浪微博诉脉脉不正当竞争案》，载《上海大学学报(社会科学版)》2017年第6期。

(85)See hiQ Labs, Inc. v. LinkedIn Corp., 938 F. 3d 985 (2019).

(86)崔国斌：《大数据有限排他权的基础理论》，载《法学研究》2019年第5期。

(87)参见崔国斌：《公开数据集合法律保护的客体要件》，载《知识产权》2022年第4期；张新宝：《产权结构性分置下的数据权利配置》，载《环球法律评论》2023年第4期。

(88)参见崔国斌：《公开数据集合法律保护的客体要件》，载《知识产权》2022年第4期。遗憾的是，很多探讨统一数据产权立法的论文还没有考虑如何界定受保护数据集合的客体范围问题。