

# 基础数据与增值数据的二元划分

## The Binary Division of Basic Data and Value-added Data

丁道勤

DING Dao-qin

**【摘要】** 数据是一种重要的资产，数据权属是数据利用和流通及数据产业化的逻辑起点，数据资产所有权的归属决定着数据价值利益的分配以及对数据质量、安全责任的划分。为了促进数据产业发展，保护数据主体合法权益，应当建立数据二元所有权，将数据分为基础数据和增值数据，两者对应不同的权属。基础数据是最本源的数据，即所有足以对主体构成识别的数据均为基础数据。用户作为个人数据的提供者，拥有个人基础数据的所有权，这是数据资产权属的基本原则。增值数据是指数据处理者对网络用户从事各种活动进行搜集整理等增值处理行为产生的各种数据，例如搜索引擎记录、电子商务记录、用户使用习惯、潜在用户群等。数据处理者享有经个人数据主体同意基于基础数据进行加工编辑分析而产生的增值数据所有权。

**【关键词】** 基础数据 增值数据 数据权属

**【中图分类号】** DF49 **【文献标识码】** A **【文章编号】** 2095-9206(2017)02-0005-06

**Abstract:** Data is an important property. Data ownership is a logical start point for the use, distribution and commercialization of data. The ownership of data property will determine the distribution of data value and the division of responsibilities for data quality and data security. In order to facilitate the development of the data industry and protect the legitimate rights and interests of data owners, the dual ownership of data should be established by dividing data into basic data and value-added data, each corresponding to a different owner. Basic data is the original data, that is to say, data which can help identify the original data provider is all basic data. As providers of personal data, users enjoy the ownership of personal basic data. That is the basic principle of data property ownership. The value-added data refers to all kinds of data generated from the value-added behavior of data processors through their collection and organization of the information on Internet user activities, such as search engine records, e-commerce records, users' habits and potential user groups. Data processors enjoy the ownership of value-added data generated from the processing, editing and analysis of basic data allowed to be conducted with the consent of personal data owners.

**Key words:** Basic data Value-added data Data ownership

**【收稿日期】** 2016-12-30

**【作者简介】** 丁道勤，男，1977年5月生，京东集团法务部政策研究总监，中国互联网协会互联网法治工作委员会委员，中国政法大学法学博士后，主要研究方向为数据保护、电子商务、网络信息安全、互联网竞争等领域法律政策。

本文仅为作者个人观点，不代表任何机构立场。

## 一、问题的提出

以下两则非常有意思的数据信息案例，引发了业界的广泛关注和热议。

一是 cookie 隐私权第一案。2015 年 5 月 6 日，南京市中级人民法院对“北京百度网讯科技公司与朱烨隐私权纠纷案”做出终审判决，撤销南京市鼓楼区人民法院一审判决，认定“百度的个性化推荐行为不构成侵犯朱烨的隐私权”，判决驳回原告朱烨的全部诉讼请求。针对同一事实，一审判决“隐私权侵权”，二审认定“不侵权”，两级法院得出了截然相反的法律判决。<sup>〔1〕</sup>

二是“脉脉”非法抓取使用“新浪微博”用户信息被判不正当竞争案。2016 年 4 月 26 日，北京市海淀区人民法院审结了微梦创科公司诉淘友天下技术公司、淘友天下科技发展公司不正当竞争纠纷案，法院认定二被告非法抓取、使用新浪微博用户信息等行为构成不正当竞争，判决停止不正当竞争行为，赔偿经济损失 200 万元等。<sup>〔2〕</sup>

两则案例涉及诸多行业政策和法律问题，其中一个共性问题是：用户使用互联网公司提供的网络信息服务所产生的数据信息，抑或对用户数据信息进行匿名化处理后产生新的数据信息，其权属如何界定，到底应当归谁所有。进而言之，在大数据时代，应当构建起何种数据权属规则来保护用户权益，顺应行业发展。

## 二、基础数据与增值数据的二元划分

正如全国人大财经委员会副主任委员、央行原副行长吴晓灵在“第一届中国金融科技大会 2016”上所指出的，数据资源与土地、劳动力、资本等生产要素一样，已成为促进经济增长和社会发展的基本要素。数据能带来价值，是一种资产，明晰产权是建立数据流通规则和秩序的前提条件。<sup>〔3〕</sup>个人数据<sup>〔4〕</sup>兼具人格权与财产权双重属性，也同时具有价值和使用价值，数据权属是数据利用和流通及数据产业化的逻辑起点。数据资产所有权的归属决定着数据价值利益的分配以及对数据质量、安全责任的划分。现有立法没有明确规定数据资产所有权的归属，数据主体与数据处理者（数据控制者）<sup>〔5〕</sup>对此存在争议。有观点认为，个人对数据拥有绝对的财产权，个人产生的数据理所当然地应当属于个人所有。无论数据如何被流转或利用而产生出的新数据，其所有权都属于个人

〔1〕 参见江苏省南京市中级人民法院：“上诉人北京百度网讯科技有限公司与被上诉人朱烨隐私权纠纷一案的民事判决书”〔（2014）宁民终字第 5028 号〕，载 <http://openlaw.cn/judgement/cff0e96c32594414bb7cce952e3f064f?keyword=%E6%9C%B1%E7%83%A8%E9%9A%90%E7%A7%81%E6%9D%83%E7%BA%A0%E7%BA%B7>，最后访问时间：2016 年 10 月 8 日。

〔2〕 参见 [http://weibo.com/3927469685/DsKHEbnwm?type=comment#\\_rnd1476064772299](http://weibo.com/3927469685/DsKHEbnwm?type=comment#_rnd1476064772299)，最后访问时间：2016 年 10 月 10 日。

〔3〕 参见吴晓灵：“谁的数据谁做主”，载中国证券网，2016 年 7 月 10 日，[http://news.cnstock.com/news/sns\\_bwkx/201607/3838818.htm](http://news.cnstock.com/news/sns_bwkx/201607/3838818.htm)，最后访问时间：2016 年 10 月 10 日。

〔4〕 一般而言，信息和资料都是数据的内容，数据是信息或资料的表现形式，信息或资料的表现形式更为多样。从相关国家和地区的立法例来看，个人信息和个人数据的保护本质上并无大的差别。为了行文方便，本文不区分“个人数据”与“个人信息”、“个人资料”等概念。

〔5〕 本文所称的数据处理者的概念也包含数据控制者。

所有。另有观点认为，用户使用互联网企业所提供的服务所产生的数据属于企业所有，应当是“谁投资，谁所有”，“谁产生，谁所有”，“谁记录，谁所有”。进而有学者认为，被记录方单方信息数据，所有权归数据被记录者。<sup>〔6〕</sup>

笔者认为，上述两种观点并非完全对立而不可调和，应当对数据权属进行二元结构化分析。换言之，应当建立数据二元所有权，不同类型的数据确立不同的所有权。借鉴《电信条例》将电信业务区分为基础电信业务和增值电信业务，数据可以区分为基础数据和增值数据，两者对应不同的权属。不仅仅是数据因被数据处理者所“记录”<sup>〔7〕</sup>或“衍生”<sup>〔8〕</sup>产生了新的所有权，关键是数据处理者对基础数据的处理应用产生了增值的效益，因而享有对增值数据的所有权。

### （一）基础数据

基础数据是最本源的数据，主要是指个人数据，即所有足以对主体构成识别的数据。根据国际通行定义，所有能够直接或间接识别特定个人身份的数据都是个人数据。例如，2016年欧盟《通用数据保护条例》第4条规定，“个人数据是指关于已识别出，或可被识别出的自然人的任何信息”。<sup>〔9〕</sup>

#### 1. 基础数据的权属

个人数据已成为用户个人资产的重要构成要素，用户作为个人数据的提供者，拥有个人基础数据的所有权，这是数据资产权属的基本原则。因为随着互联网业务应用的不断深化，个人数据的价值早已超出识别数据所有人的简单功能，从其被赋予流通功能并和现实经济活动结合在一起开始，个人数据便具有经济资源的天然禀赋，成为个人资产的重要构成要素，并演变为企业数据、行业数据、政府数据和国家数据等各种形式的数据资产。

#### 2. 基础数据的范围

在大数据应用过程中，数据的随时记录和系统化结构化处理逐渐成为一种常态，“数据化一切”已成为一种现实，个人数据被不断挖掘使用，使得与数据主体相关的数据数量和范围激增，在强大的数据收集、储存和整合技术能力下，让特定数据信息对于具体数据主体的身份间接识别几率大增。在大数据时代，个人数据数量和种类不断膨胀，数据主体很难准确区分各类数据的范围边界。若仍以既有标准界定个人信息，将造

---

〔6〕 参见陈筱贞：“大数据权属的类型化分析——大数据产业的逻辑起点”，《法制与经济》2016年第3期，第46页。

〔7〕 最高人民法院《关于审理著作权民事纠纷案件适用法律若干问题的解释》第14条规定：“当事人合意以特定人物经历为题材完成的自传体作品，当事人对著作权权属有约定的，依其约定；没有约定的，著作权归该特定人物享有，执笔人或整理人对作品完成付出劳动的，可以向其支付适当的报酬。”在没有约定的情况下，自传体作品著作权属于传主所有，并不属于“记录”的执笔人或整理人。

〔8〕 有学者从数据内容产生方式的角度，将数据分为原生数据和衍生数据（也有划分为原始数据和派生数据），能够建立知识产权的数据是衍生数据。参见杨立新、陈小江：“衍生数据是数据专有权的客体值的数据”，载《中国社会科学报》2016年7月13日，转引自 [http://ex.cssn.cn/bk/bkpd\\_qklm/bkpd\\_qkszh/201607/t20160713\\_3119714.shtml](http://ex.cssn.cn/bk/bkpd_qklm/bkpd_qkszh/201607/t20160713_3119714.shtml)，最后访问时间：2016年9月5日。

〔9〕 “personal data” means any information relating to an identified or identifiable natural person (“data subject”).

成个人信息保护范围的过度膨胀，提高各方主体的守法成本，不利于信息资源的有效利用。<sup>[10]</sup>因此，需要采取适当标准，合理界定基础数据的范围。

在国内外既有立法中，关联性、可识别性和系统化处理是界定基础数据保护范围的核心要素。但是，基础数据定义的宽泛性意味着数据经常会被推定为“个人的”，除非能够清楚地说明不可能将数据与可识别的人联系起来（即除非数据是真正匿名的）。<sup>[11]</sup>对于基础数据的判定标准，《2013年世界经济论坛报告：开启个人数据价值》提出了一些可资借鉴的判定要素，例如，数据的类型、数据涉及的实体、服务提供商的信任水平、收集方法、设备环境、应用和使用以及交易各方之间的价值交换。<sup>[12]</sup>

笔者认为，基础数据范围的判定标准可以采取“合理识别”原则，即正常理性人采取通常的方式方法可以识别的特定身份的数据构成基础数据，反之则不属于基础数据。

### 3. 基础数据的保护原则

一般数据保护原则主要包括知情同意原则、自主可控原则、透明度原则、匿名原则、最小化原则等方面，<sup>[13]</sup>基础数据保护也遵循这些原则，基础数据的总体保护原则可概括为权利保护和防止滥用原则，即应当更多地从权利保护的角度，确认包括隐私权、被遗忘权等个人数据权利是基本人格权范畴，数据主体有自我控制和排除外在侵扰的权利。

## （二）增值数据

增值数据主要是指数据处理者对网络用户从事各种活动进行搜集整理等增值处理行为产生的各种数据，例如搜索引擎记录、电子商务记录、用户使用习惯、潜在用户群等。

个人信息同时具有人格属性和财产属性，而这两种属性在社会交往过程中都可以得到充分的体现。信息的自主价值保障了个人人格的独立自主，决定了个人在社会交往中的地位优劣与否，而且这种独立性和地位一直延续到信息流转的整个过程。<sup>[14]</sup>正是个人基础数据的自由流通所产生的自主价值和使用价值，使得数据处理者利用基础数据实现数据增值。

### 1. 增值数据的权属

数据增值行为是对基础数据进行记录、检索、整理、标注、比对、分析、挖掘等匿名化处理和利用行为，即基于基础数据直接为客户创造附加价值的活动，如生产出客户

[10] 参见梅夏英、刘明：“大数据时代下的个人信息范围界定”，《中国法学》2014年专刊，第25页。

[11] 参见〔德〕克里斯托弗·库勒：《欧洲数据保护法：公司遵守与管制》（第二版），旷野、杨会永等译，法律出版社2008年版，第99页。

[12] 参见 [http://www3.weforum.org/docs/WEF\\_IT\\_UnlockingValuePersonalData\\_CollectionUsage\\_Report\\_2013.pdf](http://www3.weforum.org/docs/WEF_IT_UnlockingValuePersonalData_CollectionUsage_Report_2013.pdf)，最后访问时间：2016年10月14日。

[13] 美国《消费者隐私权利法案（草案）》规定的保护原则包括：自主控制、透明、尊重背景、安全、访问和更正、限制收集和法律责任，参见徐琦：“大数据时代美国隐私保护之困”，《中国传媒科技》2013年第9期，第43页；欧盟《数据保护指令》的原则分为数据品质要求和合法处理要求两种，前者包括公平合法原则、目的限制原则，后者包括同意、法定义务、公共利益等，参见华劼：“网络时代的隐私权”，《河北法学》2008年第6期，第10页。

[14] 参见谢远扬：“信息论视角下个人信息的价值——兼对隐私权保护模式的检讨”，《清华法学》2015年第3期，第106页。

需要的数据产品，提供客户需要的数据分析服务。数据的分析挖掘服务等数据增值行为是大数据产业的核心，是最具商业价值的行为，这种数据增值行为是在基础数据之上附加了增值性的创造活动。因为从数据的生命周期来看，可以将大数据产业分为数据的采集、整理、存储、分析挖掘和应用几个部分，应用中还包括数据的可视化。在这样的生命周期里，每个方面都会涉及相应的拥有著作权或专利权的软件、硬件开发和创新服务。<sup>〔15〕</sup>正是数据处理者在基础数据基础上实施了分析挖掘服务等创造性行为，使得这些数据产生了增值，这种增值数据权益是建立在基础数据权利人的所有权之上的。数据处理者对增值数据进行加工处理即可形成具有著作权的各种分析报告，数据处理者享有经个人数据主体同意基于基础数据进行加工编辑分析而产生的增值数据所有权。

赋予企业对于匿名化数据的所有权，有利于规范数据交易市场，遏制数据的非法黑市交易，让数据在有序可控的规则之下充分流动。<sup>〔16〕</sup>当然，承认数据处理者享有增值数据所有权，并不否认数据主体享有的基础数据所有权，例如，《微博服务使用协议》第4.7.5条就约定，“用户对微梦公司及其关联公司的前述授权并不改变用户发布内容的所有权及知识产权归属，也并不影响用户行使其对发布内容的合法权利”。<sup>〔17〕</sup>

## 2. 增值数据的范围

结合基础数据范围的界定，增值数据的范围主要包括用户使用了数据处理者的应用程序或信息服务所产生的那些不足以识别特定人身份的数据，以及数据挖掘分析产生的数据报告等。例如，搜索引擎企业所搜集的各种记录、电子商务企业所记录的用户行为数据及销售数据、用户使用习惯等方面的数据。

在文首所提的 cookie 案中，一审法院就认为，用户在互联网空间留下私人的活动轨迹展示了个人上网的偏好，反映个人的兴趣、需求等私人信息，在一定程度上标识个人基本情况和个人私有生活情况，属于个人隐私的范围。使用 cookie 技术收集了用户的网上活动轨迹，并据此展示与用户上网信息有一定关联的推广内容，进一步利用了他人隐私进行商业活动，构成侵犯他人的隐私权。但二审法院却认为，cookie 信息主要说明哪些范围的 URL（链接）是有效的。百度通过 cookie 技术收集、利用的是未能与网络用户个人身份对应识别的数据信息，该数据信息的匿名化特征不符合“个人信息”的可识别性要求。网络用户通过使用搜索引擎形成的检索关键词记录，虽然反映了网络用户的网络活动轨迹及上网偏好，具有隐私属性，但这种网络活动轨迹及上网偏好一旦与网络用户身份相分离，便无法确定具体的信息归属主体，不再属于个人信息范畴。<sup>〔18〕</sup>该案可研究的争议点很多，从基础数据和增值数据二元划分的视角来看，cookie 数据不能识别特定人身份的就属于增值数据，数据处理者可以享有 cookie 数据、检索关键词记录以及浏览网页记录等数据信息的所有权。

数据处理者一般通过应用服务协议取得基础数据的授权使用，匿名化处理后进行增值挖掘分析，这种增值处理行为遵循“合理匿名化”原则，即在当前技术条件下正常理

〔15〕 参见前注〔6〕，陈筱贞文，第46页。

〔16〕 参见王融：“关于大数据交易核心法律问题——数据所有权的探讨”，《大数据》2015年第2期，第2015018—5页。

〔17〕 参见“微博服务使用协议”，载 <http://weibo.com/signup/v5/protocol>，最后访问时间：2016年10月19日。

〔18〕 参见前注〔1〕。

性人采用通常手段进行匿名化处理后不可回溯的，应被视为履行了匿名化义务。例如，文首所提的“脉脉”非法抓取使用“新浪微博”用户信息被判不正当竞争案，法院认为，脉脉抓取用户微博的用户信息、头像、职业信息、教育信息进行展示，也没有及时删除上述信息，其行为危害到微博用户信息安全，构成不正当竞争。正如法院所指出的那样，大数据时代，保护用户信息是衡量经营者行为正当性的重要依据，也是反不正当竞争法意义上尊重消费者权益的重要内容。该不正当竞争案凸显出来的却是背后的数据之争，实质上是数据处理者基于基础数据产生的增值数据被其他主体不当抓取并利用的情形。根据《微博服务使用协议》第4.7条约定，“用户知悉、理解并同意授权微梦公司及其关联公司可非独家、可转授权地使用用户通过微博发布的内容，前述内容包括但不限于文字、图片、视频等”，第5.3条约定，“微梦公司是微博平台及微博产品中所有信息内容的所有权及知识产权权利人。前述信息内容包括但不限于程序代码、界面设计、版面框架、数据资料、账号、文字、图片、图形、图表、音频、视频等，除按照法律法规规定应由相关权利人享有权利的内容以外”。〔19〕

### 3. 增值数据的保护原则

基础数据的保护原则重在权利保护，相对应的增值数据保护原则侧重于促进数据自由流通原则。数据处理者不经数据主体同意，在整体上使用个人数据是允许的，法律所禁止的是未经主体许可披露个人数据，或将特定的个人数据挪作他用。

在信息数据化的时代，用户数据在互联网的自由流通也是常态化的事情，在一定程度上，各种数据的合理流通完全是网络开放性和分享性的必然要求，“数据不应该以它的存储而定义，应该由它的流转来定义”〔20〕。数据自由流通原则要求建立数据收集、储存、加工、转移和删除等处理环节的合理规则及匿名化和透明度规则，并推动数据资产的产权交易及程序公正。

## 三、结语

本文的结论其实就是一句话：是时候将基础数据和增值数据分开了。

数据的合法流通和利用是大数据产业发展的关键，而数据权属又是数据利用和流通及数据产业化的逻辑起点。数据资产所有权的归属决定着数据价值利益的分配以及对数据质量、安全责任的划分。近年来，欧委会正在搜集“如何规范数字所有权”的意见，探讨具体举措以帮助欧洲企业进行数字创新。消费者数据已经成为“可交易商品”，谷歌即将用户信息卖给第三方作为盈利手段。〔21〕

(下转第30页)

〔19〕 参见前注〔17〕。

〔20〕 凯文·凯利认为，个人数据才是大未来，所有生意都是数据生意，数据不应该以它的存储而定义，应该由它的流转来定义。随着云技术的不断发展，介入网络的能力要比实际拥有的所有权更重要。

〔21〕 这涉及在线活动产生的海量数据库的知识产权问题。处理这些问题仅依赖竞争政策还不够，可能的方案包括为消费者创造“个人数据空间”与想要使用其数据的各方协商如何使用等，还可设立“数据传输空间”供企业交换已对自己无用的数据。参见欧盟数字委员奥廷格：“欧盟数字委员提出数据所有权方案”，载 <http://finance.sina.com.cn/roll/2016-02-29/doc-ifxpvzah8411172.shtml>，最后访问时间：2016年10月18日。

中获取的信息，只能用于维护网络安全的需要，不得用于其他用途”，确定了政府履行网络安全保护职责所获信息用途特定的原则，所获得的相关信息只能用于维护网络安全的需要，此原则同样应适于网络安全信息共享领域。

建议进一步细化该原则性规定在网络安全信息共享领域的适用，借鉴美国 CISA 立法的相关条文，建议规定：对于通过合法共享而获得的网络安全信息，政府部门的披露、留存、处理或者使用只限于以下情形：（1）为了网络安全目的；（2）识别网络安全威胁或者网络安全漏洞；（3）应对、防止或者减轻对人民群众可能造成或已经造成的重大人身威胁和财产损失；（4）应对、调查、追诉、防止对未成年人的严重威胁，以及与国家安全、恐怖主义、电信诈骗、身份盗窃、间谍、商业秘密保护相关的特定犯罪行为。

## 参考文献

- [1] Aaron J. Burstein. Amending the ECPA to Enable a Culture of Cybersecurity Research [J]. Harvard Journal of Law & Technology, 2008.

(责任编辑：于文豪 赵建蕊)

## (上接第 10 页)

我国现有立法没有明确规定数据资产所有权的归属，笔者建议建立数据二元所有权，将数据分为基础数据和增值数据，两者确立不同的所有权。用户作为个人数据的提供者，拥有个人基础数据的所有权，这是数据资产权属的基本原则。数据处理者享有经个人数据主体同意基于基础数据进行加工编辑分析而产生的增值数据所有权，例如搜索引擎记录、电子商务记录、用户使用习惯、潜在用户群等数据信息。

## 参考文献

- [1] 吴晓灵. 谁的数据谁做主 [EB/OL]. 中国证券网, 2016-07-10 [2016-10-10]. [http://news.cnstock.com/news/sns\\_bwkx/201607/3838818.htm](http://news.cnstock.com/news/sns_bwkx/201607/3838818.htm).
- [2] 陈筱贞. 大数据权属的类型化分析——大数据产业的逻辑起点 [J]. 法制与经济, 2016(3).
- [3] 杨立新, 陈小江. 衍生数据是数据专有权的客体值的数据 [N/OL]. 中国社会科学报, 2016-07-13 [2016-09-05]. [http://ex.cssn.cn/bk/bkpd\\_qklm/bkpd\\_qkszh/201607/t20160713\\_3119714.shtml](http://ex.cssn.cn/bk/bkpd_qklm/bkpd_qkszh/201607/t20160713_3119714.shtml).
- [4] 梅夏英, 刘明. 大数据时代下的个人信息范围界定 [J]. 中国法学, 2014年专刊.
- [5] [德] 克里斯托弗·库勒. 欧洲数据保护法：公司遵守与管制 [M]. 第二版. 旷野, 杨会永等译. 北京：法律出版社, 2008: 99.
- [6] 徐琦. 大数据时代美国隐私保护之困 [J]. 中国传媒科技, 2013(9).
- [7] 华劼. 网络时代的隐私权 [J]. 河北法学, 2008(6).
- [8] 谢远扬. 信息论视角下个人信息的价值——兼对隐私权保护模式的检讨 [J]. 清华法学, 2015(3).
- [9] 王融. 关于大数据交易核心法律问题——数据所有权的探讨 [J]. 大数据, 2015(2).

(责任编辑：于文豪 赵建蕊)