

2017年8月3日

环球法律专递 GLO Legal Updates



大数据与反垄断

作者：万江

数据被认为是互联网时代的货币，大数据产业在摩尔定律的“护航”下迅猛发展，伴随而来的垄断疑虑饱受关注。2010年以来，在 Microsoft/Yahoo、Google/DoubleClick、Facebook/Whatsapp 等并购案中，反垄断机构都表达了对大数据的竞争关切，同时，反垄断机构针对 Google、Facebook、Amazon、Apple 等公司的调查无不与大数据有关。在中国，数月前顺丰与淘宝关于部分用户数据接入的纠纷也将大数据的反垄断问题推送到了中国公众的面前。

2016年5月，法德两国的反垄断机构联合发布了《竞争法与数据》的报告，OECD 竞争委员会于同年11月在巴黎召开的圆桌

会议上专门讨论了竞争法与大数据的问题。可见，大数据与反垄断是一个新问题，也是当下“数据驱动型经济”（data-driven economics）不可避免的问题。

一、什么是大数据？

数据通常是指与计算机相关的信息及信息所表征的内容。信息是多种多样的，在大数据的背景下关注的主要是个人信息。然而，大数据是什么？尽管目前已经有近二十种定义，却没有一个被普遍接受。大数据显然不仅是指大体量的数据（large volume of data），目前最广为接受的是大数据的“4V”概念，即大体量（Volume）、迅速（Velocity）、多样（Variety）和价值（Value）：

- 大体量是大数据的基本特征，如今的数据计量单位已经到 pb 级，目前 90%

的数据是过去两年中产生的，并且在未来数据量预计每两年会翻一倍。

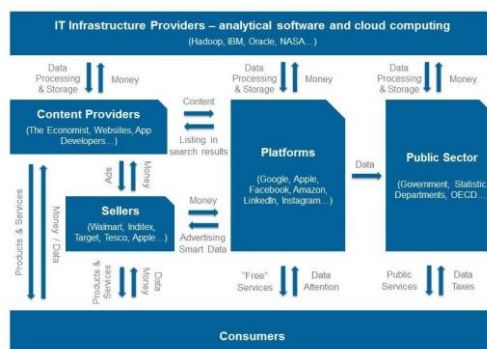
- 迅速是指数据的生产、采集、计算、传播都是迅速的，如果不能在尽可能短的时间内处理数据，数据的价值也会很快丧失。
- 多样是指数据的类别、形态、来源、采集方式等的多元化，大数据主要立足于个人数据（消费者）、结构化数据（当然非结构化或半结构化数据也是有价值的）和源于多平台的免费（自愿）提交的数据。
- 大数据的价值来自于数据的分析处理，取决于数据的采集和分析结果，依赖于复杂设计的算法以及机器的深度学习能力。

总之，大数据是数据的集合，但又不仅仅是数据，我们既看到了拥有海量数据的 Myspace 的没落，也看到了 Facebook 收购 Myspace 之后取得的巨大成功。大数据之所以具有重大价值，关键在于数据的采集、计算、分析和结果的应用。如今，人们越来越认识到大数据代表了基于海量的直接或间接相关的数据堆积，通过有效的算法进行分析，并利用分析结果向消费者提供更精确的产品或服务、甚至开发新产品或服务的一种现象。笔者较为认同

Andrea De Mauro 等（2016）给出的大数据定义：“大数据是指大体量、迅速、多样且需要特殊技术和分析手段才能产生价值的信息资产”。

二、大数据的产业生态

2015 年，OECD 在一份研究报告中对大数据的生态系统做出了非常详细的描述，我们可以通过下图简明扼要地了解大数据产业生态的要点：



在这个生态系统中，消费者处于数据提供者的角色，处于生态链的最底端，他们从内容供应商、商家和网络平台和公共部门获得各项产品和服务，同时也在供应他们的个人信息和消费数据；IT 设施供应商，也就是分析软件和云计算供应商提供数据处理和储存的服务，处于整个生态链的最顶端；而内容供应商、商家、网络平台等处于生态链的中间层，它们采集数据，委托处理数据并利用数据处理结果向消费者提供更有质量的或开发新的产品或服务。

三、大数据产业和市场的特性

大数据产业是伴随着计算机和互联网技术的发展而成长起来的，只有强大的数据存储和计算能力，大体量的数据才能迸发出巨大的商业价值，而互联网是真正实现其行业价值的最佳渠道，因为数据的采集、数据的交换、数据的应用必须要有充分的网络性和传播速度才能实现，大数据是数字市场（digital market）的重要组成部分，大数据产业是互联网产业的重要一环，因此大数据产业和市场也具有互联网的很多特性，包括双多边市场（Two- or Multi-sided markets）、网络效应（network effects）、多平台接入性（Multi-homing）和市场的易颠覆性（Market dynamics）等，有些特性更因为大数据的特征而强化。

- 网络市场的双多边性无处不在，搜索引擎（Google、百度等）、社交网络（Facebook、新浪微博等）、即时通信工具（WhatsApp、微信）、网络平台（Amazon、淘宝、携程等）都会在双向或多向面对不同的客户或消费者，且面向的每个市场都会相互影响，因此就有经营者选择在某个面向市场采取免费服务吸引客户的方式刺激或补贴另一个面向市场的拓展，这是双多边市场最大的特性。

- 网络效应是指某个用户对产品的使用会影响到这个产品对其他用户的效用，通信就是网络效应最典型的行业。网络效应与双多边市场相结合，会显现出交叉网络效应。
- 在网络世界，每个网民都会同时接入多个平台，就像每个人都可能会拥有多张银行卡一样，这也和大量网络平台采取免费接入的营业模式有关系。
- 网络市场（或数字化市场）所形成的市场力量很容易被创新型的技术和产品颠覆，这也是被那些不支持反垄断机构过多介入网络市场的人士所反复强调的。

四、大数据市场的竞争格局

在大数据市场中，内容供应商、各类商家、网络平台（搜索引擎、社交网络等）之间的竞争是最受关注的。它们在大数据背景下都处于双边甚至多边市场环境中，（1）这些不同的市场主体可能同时在两个甚至多个市场上展开竞争；（2）而且这种竞争不仅发生在内容供应商之间、商家之间、网络平台之间，内容供应商、商家、网络平台三类主体相互之间也存在几乎同样的竞争；（3）它们都有很强烈的意愿与IT基础设施供应商之间进行纵向整合，形

成产业寡头；（4）它们不仅在面对消费者提供产品或服务的市场中或者吸引广告客户的市场上存在竞争，在获得消费者信息的市场上也存在竞争，这种双、多边市场的竞争还会相互影响。

五、大数据相关市场的界定

讨论大数据市场的反垄断问题，首当其冲的是相关市场如何界定。大数据还是一个比较模糊的概念，讨论大数据定义的目的就在于挖掘大数据的一系列相关要素，大数据相关产品或服务至少可以分割为大数据包（即数据本身）、大数据的采集、大数据的传输和交换、大数据的分析和运算、大数据分析结果的运用等。正如上文所谈到的，讨论大数据的问题不能仅仅停留在“数据”上，必须同时考虑数据之外的行为和结果。两个市场主体持有的零碎数据可能会有重合，但是结构化的大数据很可能是不同的，大数据的算法和分析路径也很可能是不可相互替代的，基于不同结构的大数据通过不同的算法得出的结论却又可能是可替代的，不过，结论应用的方向又可能是不同的。我们可以看到目前市场上的搜索引擎、即时通信工具、网购平台等所采集到的数据无论是在数据内容、数据形态、数据来源、数据目标、数据隐私保护政策等都是不同的，并不是完全可替

代的。在 Facebook/WhatsApp 并购案中，欧盟委员会重点考虑的问题是 Facebook 是否会通过此项并购取得 WhatsApp 的用户数据后对市场竞争产生损害。欧盟担心 Facebook 完成此项并购后，会提升其数据采集能力，Facebook 可以依赖从 WhatsApp 用户所采集的数据提高其在社交网络平台上广告投放的精准度，从而使 Facebook 在在线广告市场等领域的市场力量得到加强。Facebook 和 WhatsApp 的用户一定会有重合，但是两个公司的数据采集重心可能不同，抓取数据的方式和隐私政策等都会有区别，欧盟委员会着重考虑了 Facebook 可以采集到其所没有但 WhatsApp 拥有的用户的数据的情形，但 WhatsApp 和 Facebook 重合的用户中，数据也还有所不同，WhatsApp 对于用户数据的算法恐怕也是 Facebook 所看重的，基于 WhatsApp 独有的算法得出的大数据分析结论对 Facebook 而言显然也是有价值的。总之，大数据相关市场的界定，不能模糊的去理解大数据的概念，在确定相关产品或服务的过程中，应当将数据及其背后的一系列行为分割清楚，根据个案精准分析“大数据”产品，它可能是数据或数据与信息采集的结合，也可能仅仅涉及大数据算法或结果运用的市场竞争，更有可能是涉及到大数据的全要素的产品或服务，这都需要

进行个案分析。但需要明确的是，尽管数据的接入是无所不在且成本低廉的（零散数据的可替代性），这不意味着结构化的数据以及对结构化的数据进行处理和分析的过程也具有可替代性。

六、拥有数据优势就意味着市场力量吗？

在谈论市场力量的时候，提到最多的是数据优势（**data advantage**），但是从反垄断分析框架出发，评估市场主体是否具有市场支配力通常考虑三方面因素：市场壁垒（针对潜在竞争对手）、企业对上下游的控制力（针对交易对手）以及对竞争对手的排挤能力（针对实际竞争对手）。我们假设一个市场主体拥有一个强大的大数据库，会不会导致市场壁垒的出现？数据具有无所不在的特性，**Facebook/WhatsApp** 并购案之所以能够获得美国和欧盟反垄断机构的无条件批准，关键在于监管机构认为市场上有大量的主体也能够免费采集到足够多的用户数据，也就是说数据的接入门槛是很低的。然而，大数据市场具有鲜明的双多边市场特性和网络效应，市场主体采集到的数据体量越大、类型和内容越多样，都将进一步提升其采集和分析大数据的能力；另一方面，尽管数据的接入通常都是免费的，但是在另一端还有一个甚至多个竞争激烈的竞价市场，市场主体采

集大量数据并以特殊技术处理数据为的是提升其在其他市场上的竞争能力，在另一个或多个市场上的成功，会通过不同的方式补贴其在数据采集和分析处理市场上的竞争力，不同企业在数据接口的数量和质量上都会表现出竞争差异性。总之，全球性的社交网络的数据采集能力和精准度要比地域性的即时通信工具或内容提供商强大得多，这种市场壁垒体现在数据优势、投资沉淀、技术门槛甚至监管政策上，还需要特别注意到的是企业在数据采集和使用过程中采取的隐私政策和隐私保护能力是核心的质量竞争要素（有学者认为这是企业采集“免费”数据的“费用”），这些也都会构成市场壁垒，因此，不能简单的认为大数据市场不会存在市场进入壁垒。在对上下游的控制力方面，拥有强大大数据资源的市场主体无疑是拥有更强大的控制力或影响力，至少那些数据商（**data broker**）更愿意和 **Google**、**Facebook**、**Apple**、**Microsoft** 等做生意，更不用说基于网络平台开发应用程序的开发商了。在横向排挤力方面，大数据的网络效应无疑会被放大，大数据的特点就是“大”，数据不够“大”，不可能支持强大的算法，也更难产生有竞争力的大数据产品。当然，在评估市场支配力的过程中，不能忽视数字市场

的技术颠覆性，这也是制约大数据市场支配力认定的重要因素。

七、与大数据相关的垄断行为

在反垄断法的框架下看，涉及任何产品或服务的反垄断行为都可以区分为经营者集中（即并购）、横向或纵向垄断协议（合谋行为）、滥用市场支配地位（剥削行为或排挤行为）等三大类。

- **经营者集中申报。**根据 OECD2015 年的报告，与数据有关的并购 2008 年是 55 件，到 2012 年就增长到了 164 件。在大数据时代，为了能够获得更多更好的数据，企业的优先策略就是并购一个拥有更大数据库的公司。目前，关于并购案件的反垄断监管讨论较多的问题包括：（1）并购申报门槛标准。当前全世界绝大部分反垄断机构所采取的都是营业额或是资产额标准，然而很多大数据公司的营业额都达不到申报的标准，但是从大数据相关市场看它们又很可能已经是一个在场巨头（incumbent giant），例如 Facebook/WhatsApp 并购案中后者的营业额就没有达到相关国家的申报门槛。目前，包括美国、欧盟、墨西哥等都在考虑降低包括大数据行业在内的特殊行业的营业额（或资产额）的

并购申报标准，德国 2017 年《反限制竞争法》关于并购申报门槛的修订也主要针对此类情形；（2）与大数据有关的并购行为，往往都是跨界并购，如 Google/DoubleClick、Facebook/WhatsApp 等，此类合并究竟属于横向合并还是跨行业合并呢？如果考虑产品因素，当然可以认为是跨行业并购，但是如果考虑其中的数据因素，则具有横向合并的性质，作为反垄断机构而言，对待横向合并和跨行业合并的态度是完全不同的；

（3）在分析竞争效果的过程中，如何看待数据采集和使用过程中的消费者隐私保护问题？这个问题存在争议，在 Google/DoubleClick 并购案中美国反垄断机构第一次在竞争分析中提及了消费者的隐私政策问题。欧盟委员会在 Facebook/WhatsApp 案中也考虑了私人数据的保障问题，但认为本案涉及的问题仍处于欧盟数据保护规则的范畴，还没有进入欧盟竞争法的范畴。认为隐私保护 / 数据保护也可以纳入反垄断法的考量范畴的理论假设是，采集或使用个人数据的过程通常都是免费的，虽然消费者也在同时免费享受某些服务，但是他们也让渡了部分隐私权利，这也可以看作是消费

者支付的隐形“费用”，而数据的相对集中可能会提高消费者的这种隐形费用成本。

- **合谋行为。**2015年，美国司法部查处了亚马逊的网商协商固定海报价格的案件，即数个网商采取一项反映消费者偏好的定价算法，分享定价信息并实施了协调价格的行为，这成为与大数据直接相关的第一起合谋案例。从理论上说，企业基于大数据而实施的合谋行为可能有：（1）企业可以利用实时数据监控一项卡特尔的实施；（2）企业可以通过分享特定的定价算法协调定价；（3）大数据有助于提高市场透明度或增进企业之间经营行为的相互依赖性，这有利于合谋行为的发生；（4）企业可能利用人工智能创造利润最大化的定价算法，通过深度学习后，造成企业之间的默契共谋，而这种基于人工智能实现的默契共谋显然是当下已有的反垄断工具无法应付的。
- **滥用大数据优势的行为。**所谓的大数据优势，包括数据数量和多样性优势、数据采集渠道和方式优势、数据的处理能力优势等，当然最重要的是指数据占有的优势。然而，有效的结

构化数据仍然会是不可替代的，甚至会构成“关键必要设施”。在大数据时代，既然数据被视为是“通货”，占有大量多样化数据或者拥有强大数据采集或接入平台的企业，无论是实施拒绝供应数据的行为、还是利用大数据资源实施搭售或歧视性行为，都有可能引来反垄断机构的关注甚至干预。2014年，法国竞争总局查处了Cegedim公司的歧视性数据接入案，作为市场上领先的医疗数据供应商，其拒绝向使用那些使用Euris软件的消费者提供核心数据，但却还在向其他消费者供应，因为Euris在健康领域消费者关系管理软件的上下游市场上与Cegedim公司具有竞争关系。在欧盟层面，涉及到数据接入的反垄断案件有“Bronner”案（1998年），“IMS Health”案（2004年）和“Microsoft”案（2007年）等，欧盟委员会正在调查的Google公司的多个行为中也有涉及到大数据相关的排挤性行为。从目前的执法情况看，尽管还有不少人对于大数据市场的反垄断监管存有疑虑，各国反垄断机构在这一领域的行动表明监管者并没有疑虑，并且与大数据相关的滥用市场支配力的行为正是反垄断监管的重点对象。

当数据发展成为大数据，其经济价值遽然爆发，也使得围绕大数据展开的经济活动在市场上蔓延，它不仅仅极大地影响甚至改变着传统的经济领域的活动，自身也在蓬勃发展，那些仅仅将大数据看作是个人隐私或数据安全的问题的人们并没有看到这一点，而担心反垄断监管会抑制大数据市场的发展和创新的观点也是基于对反垄断法的偏见，当然，监管者也不能仅用传统的监管思维和监管态度去应对包括大数据在内的一切新经济市场，大数据、云计算、人工智能等实际上也在改变包括反垄断监管在内的一切政府行为，因此，大数据行业的从业者和企业要认真看到逐步建立和发展的本行业的竞争政策和监管策略，反垄断监管机构也要积极而不失宽容的迎接大数据时代的到来。

万江 | 合伙人 法学博士

邮箱: wanjiang@glo.com.cn

版权. 环球律师事务所保留对本文的所有权利。如需转载，请注明作者姓名、作者单位以及文章来源，并保证文章的完整性。

免责. 本文及其内容并不代表环球律师事务所对有关法律问题的法律意见，同时我们并不保证将会在载明日期之后继续对有关内容进行更新，我们不建议读者仅仅依赖于本文中的全部或部分内容进行任何决策，因此造成的后果将由行为人自行负责。如果您需要法律意见或其他专家意见，我们建议您向具有相关资格的专业人士寻求专业帮助。



微信号: [globallawoffice](https://www.glo.com.cn)