

数据要素化十问

文 ◆ 中国科学院院士 中国计算机学会理事长 梅 宏

第一问：能否以及如何将数据列为资产？

资产是指什么？是指过去交易或者事项所形成的由你实际拥有或者控制的，可以给企业带来经济利益的资源。传统资产包括无形资产，它有几个性质还是比较明确的：现实性，可控数据就来了，你怎么控？权属都界定不了；经济性没有问题，数据都有经济价值，但是不是一定能带来经济价值是不确定的，带来多少也不确定的；可计量，计量性都有问题，能不能划分出来；还有可辨识，可辨识缺少依据。

对照传统的数据资产来讲，有这几个方面还是值得深究的。如果资产化沿用传统的模式，不仅中国，国际上也没有上位法，把数据作为资产。但我们又看到发展数字经济，数据资产化就是自然需求，如果说不这样它不行，或者是必然需求，这就需要探索新的途径。这件事情是我比较困惑的一件事情，怎么去实现这个东西。

第二问：如何理解数据权属性质，如何确权？

我们怎么理解数据权属的性质，怎么样确权？从法律视角来看，它是不是可以产生可以确认的产权，可确认的收益，权益归属于谁？这么多数据，非竞争性、非排他性、非稀缺性、非耗竭性。场景

是多样的，观点是复杂的。逆势定律，一个是动力，一个是独立，动力就是你的产权，我拥有它才可能产生这个东西，实际上是失效的。数据权属的复杂，大部分都与所有权相关，怎么解决这个问题。

不同类型的权属问题，比如说个人数据一定和我本身生物特征、财产特征关联的，能不能和人格权关联等同呢？另外讲公共数据，它的资源归属范围如何界定，特别是我们国家来讲，公共数据是全民所有，企业数据是否构建新型的民事权力体系，解决企业所有，还有数据权力的指向问题，甚至还有知识产权的成分，数据产品有的涉及你的智力劳动，可以呈现你的 IP，还是说中间有个人权力，还是说可以发明一种全新的权力，现在创造了很多权，把所有权搁置。

权力还有权力的归属，是否可以构建一个多元权力并存、共享呢？我们搁置争议，先往前探索是非常好的举措。权属确认实际上是全球的难题，因为没有很好的技术方案。现在世界各国现行的全国性的法律，都没有立法，现在采用的是法案进行个案的判，涉及数据侵权的判例，基本上都是用责任规则来判，而不是财产规则。我们现在面临着如何构建适应数据要素本质的签署体系和确权机制，这个大家都在探索。

第三问：如何度量评估数据价值？

我们在座不同行业的人，可能对同一个数据的需求就是不一样的，有的人愿意花高价买，有的人你送我都不要。我们有各种评估的方式，有成本法、收益法、市场法，数据评估就不一样，落后元主体，成本不容易区分。市场法不是特别适合到目前来讲，我们从数据定价的问题来看，作为一个新兴的商品一定有它的特点，成本构成的特殊性，高生产成本，低边际成本，需求多样，估值有困难，由于不同的需求，真实性难验证，隐私权难界定，还有复制运用，偷偷用了就很难防范。商品售卖的形式是什么，套利行为怎么防范，隐私信息怎么保护，还有真实性的验证等等。本身这种东西定价是由供需双方在市场上通过一种博弈活动的行为来产生。

所以说传统的资产评估方法现在看来是难以直接用于数据资产评估的，价值评估没有统一规则，参与的主体需求各异，也很难形成统一的价格标准。目前，会计核算方法还没有办法把数据估值定价，为数据估值定价提供依据，就算想入账，怎么入，会计的核算方法都不支持。

第四问：什么是数据要素的基本度量单位？

如果数据作为要素交易，它的基本度量单位是什么？

传统的要素，土地按平方米、按亩，总归有一个说法，劳动力的时间、货币有统一的度量。只有技术和企业家能力这两块比较难度量，通常的要素里面会不会谈到，它作为生产要素的促进者，是生

产效益提升角度，它不是直接按照这个去度量的。现在数据的问题来了，我们按数据本身的规模，就是基本单元，一个G、一个T，一个T就肯定比一个G值钱吗？肯定不是这样。如果按表还是按照操作系统的一个文件来，都是不好说的，因为这是两个维度的事情。

我前天在论坛上谈到数联网，用DOA体系，把现实物理数据数字化变成一个对象以后，我能不能用它来度量，这是我的一个问题，能不能做到进一步的尝试。所谓DOA把现实数据和资源进行建模，变成数字对象，然后联网变成数联网，也许是一个做法，至少没有度量单位。所以说度量单位和价值是正交的二维，属于形式或者技术维度，但是你的数据给我的形式是一张表还是什么。民间的黑市交易很清楚了，按照行来数，偷偷卖个人信息的、手机数据的、家庭数据的，那是黑市交易，那是按照条数来算的。其他还有什么办法？数据价值我理解在很大程度上取决于它所表征的物理世界对象，在物理延伸的数据是有的，但是如果没有表征物理世界的基础数据延伸不了，也许我们把DOA作为物理数字对象，作为数字孪生能不能作为度量单元，我是想找几个人琢磨琢磨，开展这个研究来做这个事。

第五问：如何构建高效数据流通交易体系？

我们讲数据交易，以数据产品和数据交易目标的交易行为，广义范畴下可能涉及数据价值深度挖掘等数据交易产业链相关交易与技术服务，这是广义的交易的定义，但我是认可的。我们现在谈的数

据交易模式，基本上全世界都是这样，一个是平台。我们有一些交易参与的平台，供给方和需求方的交易，包括三方、四方各种模式都有。还有直接交易，比如说需方和供方的直接交易，黑市交易就属于直接交易，还有一种数据包，我们两个商量好了，我一包数据给你，这是没有问题的。

还有 API，很多平台企业 API 经济已经成为数据交易里面非常活跃的成分，民间已经产生了。国际上的交易平台已经有很多，但是国际的平台基本上还是企业在做的，类型也是多样的，有很多也涉及服务。国内的平台从 2014 年构建开始到现在已经有 50 多家数据交易所，大多数是政府主导的。这里面就有一个问题，我们总是喜欢用交易所这个概念来命名这些东西。交易所是工业社会的产物，我去查了交易所的定义，是交易某种信息或者物品的信息平台或者场所。这基本的特点：交易物权清晰，交易所属于抽佣中介性质，准入制和市场化结合。

比如说，传统交易所主要是证券交易和商品大宗交易，可以是现货也可以是期货。世界上第一个交易所是阿姆斯特丹的一个交易所，数据交易一定要和信息服务区分开，数据交易是不同主体之间，针对数据不管是有偿的、无偿的一种交换行为，交换对象可以是原始数据也可以是加工后的数据产品。信息服务是在已有的信息数据基础之上分析之后形成的服务产品，他给你直接提供信息、提供支持，服务方要承担责任的，如果提供错了，这是我们交易所不能解决的。现行的交易所，各地的探索是很好的尝试，但是也面临着在权属不清晰的情况之下

规则完善的问题，交易所到底能不能交易，就是交易使用权，交易加工以后的东西，如果说交易所也提供数据的信息服务，可能就会带来问题——交易所要承担责任。你提供的服务质量不够那我就找你，你就变成一个数据信息的服务商。

国内有 50 多个交易所了，中国的特点我一直在探讨，中央 20 条明确数量要控制，交易所会不会成为各地流通的障碍。当地政府一声令下，本地企业必须本地交易，全国市场怎么建？显然是不符合，现在数据还要追求全球交易、跨境交易，这肯定是不符合。还有的交易所是政府主导的，我以为还是要允许场外交易，怎么去规范，怎么去兼容市场化的交易方式，这是多元的交易体制还是需要的，这个东西要有效还要高效，现在我们要积极地探索。

第六问：如何合理分配数据收益？

价值的利益分配，需要在保障数据基本权利的情况下使各方的收益最大化，如果纯粹靠市场机制现在似乎是有问题的。前几年靠市场机制实现这种东西，我们每一个个体的数据，贡献的每一个平台，虽然个人是没有收益的，但是我得到了服务，换取了方便。市场初期规则不完善，监管也缺失，前几年政府没有监管的领域，所以现有的收益分配明显有太多的不合理之处。分配，就是政府介入监管之后的分配怎么分，其实已经有很多的研究实践，比如以欧洲为代表的数字税就是非常重要的一个方向。

到 2021 年 10 月，G20 和联合国世界经合组织 136 个辖区已经对数字经济税收挑战形成了一种共

识，但是好像明确开始收税的还没有。这里我也提出一个问题，数据税是国家收的，收完之后能不能够普惠到每个个体，有没有一种方式探讨所谓的数据费，是指个人的数据能够获取。现在我们谈到技术发展趋势有一个 3.0 的出现，强调我的数据我作主。每个人有自己的数据空间，谁用我的数据空间，你象征性地给我一分钱也行，也算是我的收益。现在国际上已经有人在做，个人数据汇聚到中间，作为信托作用。

我觉得还是要充分考虑在未来的世界，每个人的数据能不能获益。因为毕竟是未来数字世界最基本的数据的贡献者，我们应该考虑，特别是我们国家更应该考虑。再一个是公共数据，是全民所有特性，他能不能参与利益分配，我一直存疑的。谁代表他去参与分配，如果参与分配怎么体现公平性和公正性，公共数据面临的一个难题。所以按照 20 条的讲法，体现原则，促进公平，这是没有问题的，这是我们追求的基本原则。

现在也在追求多元化数据要素的数据分配，市场评价贡献、按贡献决定报酬的机制，这件事情说起来容易，后边还是很难的。还有政府对要素收益进行调节机制，这件事情目前也没有章法，这两件事情各自都没有解决，再加到一块儿，具体的路径在哪里，我们还需要加紧探索，哪怕是试错也要探索。

第七问：如何实现公共数据的真正开放？

所谓的数据开放，就是提供方无偿地提供数据，单项提供数据，没有其他的介入。国内数据开

放，我们从 2022 年开始截至去年 10 月，有 208 个省级和城市地方政府上线了开放平台，我们 330 个地级行政区建成了公共数据开放平台，已经占了全国的 58%。国际上也在开放，开放的都是公共数据，国外的数据开放主要的模式是政府主导的、民众主导的、行业协会方主导的。2009 年美国是第一个全球开放政府数据门户网站，涉及农业、制造、天气、教育、能源等方面。

2023 年 5 月已聚集 25 万多个数据集。DATA.GOV 的最大特点是开放政府网站的应用接口 API，并统一数据格式。我们国家也在做，但是公共数据的质量和数量仍有不足，还需要进一步的提升。我一直觉得公共数据的真正开放应该是原始数据的开放，现在 20 条规定是可用不可建，是授权开放。

不管怎么样已经很好地走出一步了，我觉得挺好，但还是应该是原始数据开放。为什么这么说？公共数据取之于民，理应用之于民，我们把敏感数据剔除掉以后，来进行运用。很多数据应用的创新绝不是坐在办公室想出来的，是民间需求刺激。我们看到很多民间依托数据去创新运用的例子，还有很长的路要走。

第八问：如何平衡发展与安全？

安全和隐私问题非常突出，数据泄露事件频发，网络诈骗行为猖獗，还有数据的滥用、误用也严重。

我们国家这几年一直在加强数据安全的法律法规建设，数据安全，隐私保护，还有法律、行政法规和部门规章，这几者非常重要。按照总书记在网

信会议上的讲话，安全是发展的前提，发展是安全的保障，发展是最大的安全。现在数据应用需要数据要素化这件事情，一定要强化安全意识和底线意识，但是我们不应该设置过低的安全底线，没有绝对安全，所以你的安全底线设置过低了，那就别做，不用可能是最好的安全。因此我一直的观点是发展先行，安全跟进，用发展保安全。毕竟这一轮信息化是我们中华民族千载难逢的机遇，所以我们一定要抓住。

第九问：如何为数据要素化提供技术支持？

发展数字经济离不开数字基础设施，也离不开物理基础设施的数字化，离不开信息技术，特别是涉及数据的全生命周期的大数据技术的支撑。我把大数据技术分成四大板块：管理、处理、分析和治理。我这里列了大数据五个方面的挑战，以数据为中心的计算体系怎么构建，满足大数据高效处理，实现数据和应用的完全解耦支撑跨越的数据管理，实现多源异构大数据的多模态融合分析与可解释性分析，还有如何形成系统化大数据治理框架与关键技术。

尤其是第五个，是治理框架和治理体系，这件事情我们在研究，从基本理论到框架，还有贵州实践。怎么支撑数据的汇聚融合，技术上怎么去保障它的质量，技术上的隐私和安全保护，怎么去促进开放融通，怎么构建相应的技术标准，构建相应的生态，这些都是技术问题。所以数据要素化是一项系统工程，制度建设是最关键的东西，但是同时也是一项技术性很强的工作，需要高效有效的技术提

供支撑。因此数据治理相关技术的研发将是大数据技术体系发展的重要方向之一，我们一切要围绕数据要素化整个过程提供相应的有效的技术支持。

第十问：数据要素如何加入生产函数？

传统工业世界是讲生产函数，我就想知道数据作为要素加入之后，生产函数的位置是什么？所谓的生产函数是描述了一定技术条件下投入与产出之间的关系，即生产要素投入量和产品产出量之间的函数关系，体现了企业将一定的投入转变成产出的能力。通常来讲，土地、资本、劳动力的投入是直接变量，而通过把企业家的才能和技术作为效率提升的工具，同样的东西给一个有名的企业家和给一个新手可能产生的结果就不一样，所以技术和管理是作为效率提升的东西。

我看了很多短期生产函数、长期生产函数各种各样的东西，我们还要看到数字技术发展之后带来的一个很重要的东西，很多事情边际成本趋为零，特别是很多技术如果实现了把边际成本转化成固定成本的时候，就会带来很重要的产业拐点。传统的生产函数把劳动、资本、土地的要素作为有形的可以量化的投入，属于函数的自变量，而技术企业家不列入里面。数据好像两边都有，它作为物理世界因为对应是可以直接投，但是数据有技术含量的时候，它也会提升这个东西。所以我想看看哪些经济学家能够在生产函数里面把数据的位置描述出来，我们找到实际的案例做这件事情。■

（文章来源：工信头条公众号）