数据产品生产需要对数据制造商进行分级分类

Hierarchical classification of data manufacturers for data product production



叶雅珍（1985– ），女，博士，复旦大学计算与智能创新学院教师、上海市数据跨境应用基础研究创新中心副主任、上海市数据科学重点实验室数据资产研究室主任、上海数据研究院特聘研究员、浙江省数据知识产权研究基地智库专家，主要研究方向为数据科学、数字经济、数据资产、数据知识产权、数据跨境、数据流通和数字化转型。



朱扬勇（1963– ），男，复旦大学教授，上海数据研究院学术副院长，上海市数据科学重点实验室创始主任，《大数据》期刊编委会副主任，数据科学家。主要研究方向为数据科学与数字经济。在数据领域做出开创性工作，提出数据自治、数据财政、数据资产等概念和体系。兼任数家政府机构顾问，主导编制了多个地方政府数据发展规划。发表论文200余篇，持有高价值专利15项，出版学术著作17部。

中图分类号：TP399　　文献标志码：A　　doi：10.11959/j.issn.2096-0271.BDR26398

医疗健康数据被认为是应用价值最高的一类数据。近年来，医疗数据进入市场交易的案例持续出现，例如，山东第一医科大学第一附属医院完成了“肝病临床谱系与移植状态数据集”的数据交易；北京同仁医院与恒瑞医药、拜耳医药达成协议，实现“眼科全维度随访数据集”交易，等等。随之而来的问题是：医院是否具备数据产品开发能力？医院是否能够销售数据（营业执照经营范围是否覆盖数据经营活动）？医院能够开具数据产品销售发票吗？答案显然是否定的。医院属于公共服务机构，核心业务是医疗卫生服务，其法定职能与经营范围均未覆盖数据产品生产、销售等商业化经营活动。

除了医院，其他数据持有主体是否具备对其所持有的数据进行开发、经营的资质呢？答案基本也是否定的。医疗数据可能涉及国家数据安全及个人信息保护相关问题，无序的数据供给与数据产品开发将带来风险。因此，需要厘清数据原料供给、数据产品生产、数据商品流通、数据产品使用四大环节中各市场主体对应的能力要求、责任要求和利益分配规则。

1 数据产品（服务）开发应用的运行模式

数据产品（服务）开发应用由原料供给、产品生产、商品流通、产品使用四大环节构成，如**图1**所示。

- 在原料供给环节，各类供给方的数据库及多元数据源通过数据要素标的化机构，将所持有的数据输出至数据要素市场，其中公共数据是其主要来源。
- 在产品生产环节，数据制造商从数据要素市场购入生产所需的数据要素，运用相关技术将形态多样的大规模数据资源（数据原料）加工形成标准化的数据产品

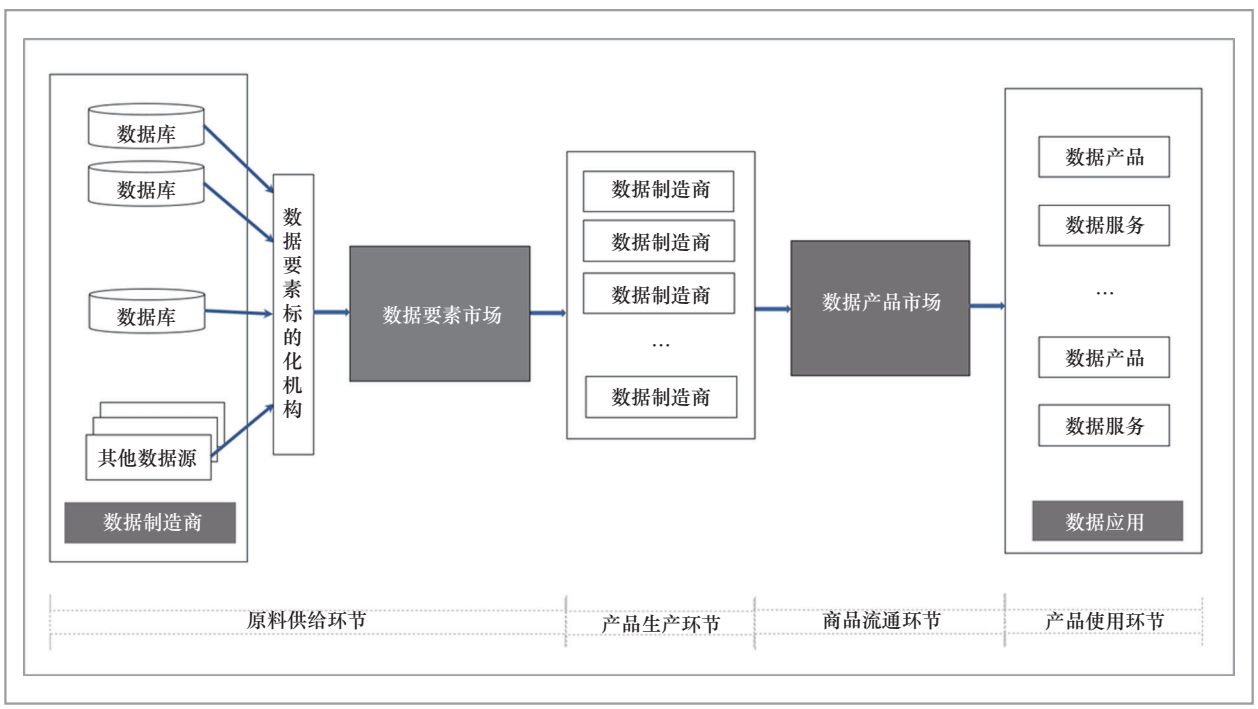

图1 数据产品（服务）开发应用流程

（服务）。

• 在商品流通环节，标准化数据产品（服务）被投放至数据产品市场，面向社会不特定主体开展市场化流通与公开交易。

• 在产品使用环节，从数据产品市场采购的数据产品（服务）可赋能千行百业，广泛应用于多元终端场景，为各行业数字化转型与智能决策提供支撑，充分释放数据价值。

在上述环节中，数据要素市场与数据产品市场作为两大市场类型，在整个数据产品（服务）开发应用过程中起着重要的衔接贯通作用。数据要素市场主要向数据制造商供应数据要素，对买方设有明确的准入限制，以此规范数据要素采购行为、明确数据使用边界与安全保护责任，杜绝不合格主体入场，阻断数据违法流转风险。数据产品市场面向所有数据需求方（数据使用者），为其提供各类标准化数据产品（服务），进入市场流通的数据产品（服务）必须通过合规认证并取得合规证书，该市场原则上不对买方设置准入门槛。

2 数据制造商资质思考

数据制造商是以数据产品生产为核心业务的新型组织形态，在数据产品（服务）开发应用中发挥着关键作用。

以瓶装水生产管理为例：瓶装水直接关系公众健康与安全，并不是任何市场主体都能开展生产经营。瓶装水制造商须依法取得政府主管部门颁发的生产许可证，并依据国家相关标准组织生产。一旦瓶装水出现饮用安全问题，相关制造企业将被依法追责；视情节严重程度，可吊销其生产许可证，取消其生产资格。

与之类似，数据产品涉及国家安全、商业秘密和个人隐私，因此数据产品的生产制造同样应该实行严格的许可证制度和产品国家标准。由于数据产品门类多样、结构复杂，需要针对不同数据产品设计设置不同的生产许可要求，这是一项长期系统性工作。

当前国家将数据分为普通数据、重要数据和核心数据三大类别，基于此，现阶段可先行考虑将数据制造商的资质划分为以下3类。

• 普通数据资质：具备数据产品生产制造能力，仅可采购普通数据；配备一定数量的认证数据工程师，建立基本的生产管理和质量控制体系，能够按照国家数据产品标准组织生产，具备配套售后服务与追溯机制。

• 重要数据资质：满足普通数据资质全部要求，且具备较高水平的数据安全保护能力；可采购重要数据，为特定用户（如部分政府部门）生产制造数据产品；也可以通过脱敏、脱密等安全处置措施，确保生产制造的数据产品不再包含影响国家数据安全的内容及个人敏感信息，以此向公开市场供给普通数据产品。

• 核心数据资质：满足重要数据资质全部要求，同时具备更高等级的数据安全保护能力；可采购核心数据，为特定用户（如部分政府部门）生产制造数据产品；也可以通过脱敏、脱密等安全处置措施，确保生产制造的数据产品不再包含影响国家数据安全的内容及个人敏感信息，以此向公开市场供给普通数据产品。

3 建立风险利益共担机制

现阶段，数据持有者的供数意愿总体不足，尤其是公共数据持有部门。其核心

原因在于各参与主体的责任与利益不匹配。在责任划分方面，当前数据安全责任主要按照谁所有谁负责、谁持有谁负责、谁管理谁负责、谁处理谁负责、谁使用谁负责的原则确定。在利益分配方面，数据流通的获益方主要是数据处理者与数据使用者。在这一框架下，数据所有者、持有者和管理者承担了主要的数据安全责任，却难以获得对应的收益，由此催生“不愿供数、不敢供数”的现象，进而导致“供得出、流得动、用得好、保安全”的方针在数据“供得出”环节就遇到了难以逾越的障碍，使得数据产品（服务）开发应用无法推进。

为破解上述困局，应当确立“谁获益、谁担责”的风险利益共担机制，推动责任主体与实际获益者相统一，明确数据所有者、持有者、管理者、处理者、使用者等的权责边界和利益分配。当数据所有者、持有者、管理者、处理者、使用者等均为市场主体时，市场机制会自发调节各方责任和利益的平衡。针对公共数据，其所有者、持有者、管理者为政府公共部门，现行制度框架并不支持这些部门获取数据流通带来的经济利益，或其所要承担的责任风险往往远超于可获得的收益。因此，有必要合理减轻这些部门的数据安全责任，将相关责任转移到数据获益主体。

分级分类设立数据制造商资质，由数据制造商依据其所持有的资质等级，确定可采购的数据原料类型及可加工的数据产品类型，并承担相应的数据安全责任。这样做既符合数据开发应用的客观实际，又符合“谁获益、谁担责”的风险利益共担原则。

从实践层面来看，即便是核心数据，一般也由专业化市场主体开展开发工作，而不是由数据持有部门自行开发，数据持有者通常不具备数据产品开发能力。在传统信息系统建设领域，通常由具有不同信息安全保密资质的市场主体承建具有安全保密要求的信息系统。同理，涉及安全保密要求的数据开发，也应由具有安全保密资质的数据产品开发主体来承担。因此，现阶段先行设置普通数据资质、重要数据资质和核心数据资质3类数据制造商资质，是符合方向性的制度安排。后续随着数据产品开发应用的深入，还可进一步细分出更多的数据制造商资质类别。

落实“谁获益、谁担责”的风险利益共担思路，还需要建立新型利益分配机制，保障提供数据的政府公共部门能够依规获得相应收益。 □