数据编织，大数据的新风口？

本报记者 李佳师

怎么实现“数据找人而不是人找数据”的梦想？“数据编织”（Data Fabric）悄然登场。2018年“Data Fabric”首次出现在Gartner的十大数据与分析技术趋势中，以后每年它都会出现在其中。10年前大数据概念在国外兴起后，不到3年就被中国用户广泛接受，而今天大多数中国厂商则是按兵不动，这又是为何呢？

“数据编织”悄然登场

Data Fabric的中文名字到底怎么翻译，IBM公司与Gartner有了分歧。IBM大中华区科技事业部云计算与认知软件部数据与人工智能信息架构产品总监王积杰将其称为“数据经纬”，而Gartner高级研究总监孙鑫认为“数据编织”更为合适，因为他认为数据编织更凸显“动态”。

在Data Fabric出来之前，数据结构的设计都主要部署成静态基础设施，而在未来将需要采用更动态的数据网格方法全面重新设计。

孙鑫在接受记者采访时表示，Data Fabric不是一个产品而是一种设计理念，是利用AI、机器学习和数据科学的功能，访问数据或支持数据动态整合，以发现可用数据之间独特的、与业务相关的关系。

而IBM对Data Fabric的看法，与Gartner专家提到的“动态”“数据网格”和AI赋能并无冲突。IBM中国研发中心首席技术官赵军伟解释：“经纬作为名词，本意为织物的直线与横线，引申为连接万物的规律，《左传·昭公二十五年》中写道——‘礼，上下之纪，天地之经纬也。’作为地理概念，经纬度可以定位地球上任何一个位置，而‘数据经纬’则可以在纷繁复杂的企业数据目录里定位任意一个数据源。经纬用作动词，是规划治理的意思，《周书·静帝纪》中提到的‘经天纬地’就是治理天下的意思。”

“我们可以把Data Fabric想象成一张虚拟的网，这张网并不能理解为一种点对点连接，而是一种虚拟连接，每个节点都可以是不同的数据系统，不同系统上的数据在网上都可以迅速被定位和找到。Data Fabric的主要功能是把正确的数据，在正确的时间里，给到正确的人。通过Data Fabric，对的人可以从对的地点，在对的时间里，获取对的数据。”王积杰对记者说。

现在的数据连接的架构设计还主要是“人找数据”，而Data Fabric设计核心是“数据找人”，在合适的时间、将合适的数据推送给需要的人。

为什么Data Fabric将会成为一种趋势，为什么越来越多的企业将在未来采用这样的方式进行部署？王积杰谈及了数据利用结构模式的变化。传统IT时代，无论是早年的“数据仓库”还是近几年的“数据湖”和“大数据”时代，其实数据利用都是集中式的架构，把数据收集到一起，让企业的数据分析师、BI（商业智能）分析师对数据进行分析。但在云计算时代，用户业务部署在多云的环境下，要想将分布在不同云上的数据集中在一起成本很高，也很费劲，于是采用去中心化、分布式的数据网络架构就成为了必然选择。

Data Fabric可以同时给业务和技术团队带来明确的价值，王积杰表示，从业务层面来看，由于企业能更容易地获得高质量的数据，从而能更快和更精确地获得企业数据洞察。从技术层面来说，由于数据复制的次数和数量较少，从而减少了数据集成的工作，方便维护数据质量和标准，也减少了硬件架构和存储的开销。由于减少了数据复制和大大优化了数据流程，加快并简化了数据处理过程，从而通过实施自动化的整体数据策略，减少了数据访问管理的工作。

Gartner认为，随着数据的日益复杂以及数字化业务的加速发展，Data Fabric已成为支持组装式数据分析及其各种组件的基础架构。由于在技术设计上能够使用/重复使用及组合不同的数据集成方式，Data Fabric可缩短30%的集成设计时间、30%的部署时间和70%的维护时间。IBM 7月发布的Cloud Pak for Data4.0的软件组合增加了智能化的Data Fabric功能，其中AutoSQL（结构化查询语言），可以通过AI来自动访问、整合和管理数据，可以帮助客户以8倍的速度、不到一半的成本，获得分布式查询的答案。

如何“编织”数据？

要实现“数据找人而不是人找数据”，Data Fabric究竟如何“编织”？

王积杰认为，Data Fabric至少需要四个维度的能力。一是能够在数据之间建立虚拟链接，简化数据访问的模式，从而减少数据复制的数量。二是需要建立一个企业数据目录，并需要利用AI技术，自动化地实现基于语义和知识的分析，理解数据及其业务含义，并建立知识图谱，从而使数据目录变得智能化和自动化。能够让需要数据的用户，随时了解到需要的数据在哪里、数据质量如何等。三是建立自动化数据平台，允许用户通过自服务的方式，访问并获取数据。四是通过提供整体的自动化策略，确保数据安全，增加数据的隐私和权限保护，并提高数据质量。

数据编织是一种新的设计理念，它是数据管理、数据收集理念的变化，与数据仓库、数据湖等技术并不是替代的关系，既可以运用现有的数据中枢、数据湖和数据仓库的技术和技能，也可以在未来加入新的方法和工具。

孙鑫谈到了实现Data Fabric的一些关键技术，比如增强型数据目录，要想实现数据找人，而不是人找数据，需要增强的数据目录，它要涵盖用户使用数据的频度与机制，了解数据与业务的关系，还包括知识图谱，通过知识图谱找到数据与业务之间的关系，找到元数据利用的整合策略，也包括推荐引擎以及在数据准备阶段的低代码等工具，低代码工具的作用在于降低数据使用的门槛，加速数据产品化。

从Data Fabric推动的难点来看，“一是理念层面的难题，中国的用户还没有意识到，数据利用和使用的方式已经发生改变，传统的集中收集再利用的方式已经不能满足需要。二是目前很多企业对于元数据不够重视。三是从人的角度看，需要提升企业数据工程师对知识图谱、图语言、图建模等数据工具的能力培养。四是数据编织的实现并不是找到一个厂商就能够完成，它是一个旅程，需要分几步走。”孙鑫认为，从用户的角度看，率先采用Data Fabric的是金融电信行业以及数据应用场景比较复杂的用户。

在这一点上，王积杰表达了与孙鑫一致的观点——这是一个方向，但并不能一蹴而就，用户需要分步实施，关键是要意识到趋势，在后续的项目实施中，按照Data Fabric的理念来构建。

国内厂商为何按兵不动？

尽管Gartner、Forrester等分析机构在几年前就提出Data Fabric是数据利用与分析领域的革命性变革，是未来方向，但记者联系国内大数据相关领域企业进行采访时发现，对此了解或进行布局的企业并不多，甚至找不到。

这与国内大数据厂商的分布有关。“国内有很多做数据库的企业，也有很多做BI（商业智能）的企业，但做数据整合的企业很少。而事实上，在国外做数据编织的往往是数据整合、数据虚拟化的厂商，这就很好理解为什么国内的大数据厂商迟迟未入场Data Fabric，因为这类企业就不多。”孙鑫告诉记者说。

大厂商没有入场很好理解，因为在Data Fabric的理念下，往往需要采用点和边的新方式去描述数据关系，需要知识图谱、图数据库等，这往往是新锐公司在做的领域，而大厂商往往有自己的数据整合工具，他们都希望在自己的平台上进行整合，但是这样的局面一定会在之后的几年发生变化。

“Data Fabric这个概念在国际上已经热起来了，但目前国内的IT用户知道的人还不多。10年前大数据的概念在国外兴起后，不到三年就被中国用户广泛接受，未来这个Data Fabric概念，中国将需要多久接受并加以应用呢？等待时间给出答案。”王积杰说。

2021年度软件和信息技术服务企业竞争力报告发布

本报讯 8月27日，中国电子信息行业联合会主办的2021年度软件和信息技术服务企业竞争力报告及前百家企业发布会在北京举办，中国电子信息行业联合会秘书长高素梅、专家委主任董云庭、副秘书长李杰出席发布会，会议由副秘书长徐金声主持。

会上发布的《2021年度软件和信息技术服务企业竞争力报告》指出，竞争力前百家企业的主要特点如下：

一是迎难克艰，规模效益双提升。2020年，在突如其来的疫情中，以软件为代表的数字经济体现出较强的发展韧性和潜力，为稳定增长、科学抗疫发挥了积极作用。竞争力指数前百家企业（简称：百强企业）2020年软件和信息技术服务收入合计18516亿元，同比增长16.7%，高于全行业平均增速3.4个百分点。百强企业中，软件业务收入规模超过100亿元的企业有20家，入围企业软件业务收入门槛超过15亿元。百强企业实现利润总额4279亿元，同比增长31.3%，高于全行业平均增速23.5个百分点。

二是创新驱动，转型升级加快。百强企业研发投入合计3967亿元，同比增长23.4%，高于同期软件业务收入增速6.7个百分点，企业平均研发投入强度超过10%。2020年，百强企业软件著作权登记量5.4万件，获授权专利数量28万件，其中发明专利占比超过70%。百强企业一方面着力夯实产业基础，另一方面，积极培育新动能，以云计算、大数据、人工智能、5G等领域的新兴软件为牵引，不断加快发展平台软件。为构筑具有国际竞争力的产业体系，支撑新型基础设施建设，引领产业转型升级发挥了重要作用。

三是开放合作，深耕国际市场。2020年，受疫情和其他因素影响，全球化的供应链收紧，导致软件技术、产品和服务的创新合作及市场拓展难度加大。面对复杂的形势，百强企业迎难而上，继续推进产品和服务出海，深化海外市场布局。2020年百强企业软件和信息技术服务出口超过200亿美元，同比增长4.7%，高于全国增速7.1个百分点。以更高水平开放引领高质量发展，是产业发展的重要方向，更是百强企业责无旁贷的使命。

四是融合渗透，推动软件定义。百强企业引领推动“软件定义”向生产生活各领域延伸。百强企业中，有三十余家企业的业务涉及推动传统产业转型升级，助推数字化车间、智能工厂普及率逐步提升，形成涉及钢铁、能源、化工、机械、家电等领域的一批跨行业跨领域平台。同时，百强企业牵头参与培育开源社区，加快构建开源生态，持续丰富应用场景，吸引产业链各方共同参与、密切协作，合力构建了富有创造力的软件生态。此外，百强企业还积极探索新业态、新模式，引领催生了分享经济、平台经济、算法经济等新形态，壮大了一批新场景，为软件产业创造了更为广阔的市场空间。（徐 恒）

2021年4月27日，国务院正式通过了《关键信息基础设施安全保护条例》（以下简称《条例》），并于9月1日起施行。这是我国首部专门针对关键信息技术设施安全保护工作的行政法规，同时也是《中华人民共和国网络安全法》的重要配套法规。《条例》对关键信息基础设施的范围、各监管部门的职责、运营者的安全保护义务以及安全检测评估制度提出了更加具体、操作性也更强的要求，为开展关键信息基础设施的安全保护工作提供了重要的法律支撑。《条例》出台的背景及必要性是什么？对于运营者来说如何才能做到合规？

我国关键信息基础设施安全保护迈入新阶段

赛迪智库网络安全研究所所长 刘权

《条例》的出台必要且迫切

关键信息基础设施涉及公共通信和信息服务、能源、交通、水利、金融、公共服务、电子政务、国防科技工业等重要行业和领域，一旦遭到破坏、丧失功能或者数据泄露，可能严重危害国家安全、国计民生、公共利益。当前迫切需要加强关键信息基础设施安全保护力度。

一是关键信息基础设施关乎国家安全和社会稳定。随着信息化的快速普及和发展，关键信息基础设施作为事关国家安全和社会稳定的重要战略资源的地位日益凸显。首先，互联网的飞速发展，使得网络入侵和网络攻击事件频发，严重威胁着关键信息基础设施的正常运转，给国家安全带来极大隐患。其次，关键信息基础设施是恐怖主义和网络攻击的重点对象，各国均将其视为网络安全的重点并上升到国家安全的高度。通常认为，关键基础设施或关键信息基础设施是支撑国家安全和公共利益的重要基础设施。同时，国家间的网络战争威胁日益加剧。信息技术的快速发展极大地开拓了互联网络平台，网络攻击不再仅仅依附于传统的常规战争而存在，已经拓展和波及所有与网络相关的事件和人员，通过技术手段破坏关键信息基础设施从而导致通信瘫痪、基础设施停摆等，已然成为网络战争的重要手段。基于此，为了更好地应对各种形式的网络攻击，维护国家安全和社会稳定，应加强对关键信息基础设施的保护。

二是加强关键信息基础设施保护是社会持续运转的重要保障。关键信息基础设施为国家机构、各行业正常运转提供必需的支撑和服务。关键信息基础设施承载或支撑着各行业的关键核心业务，是政府部门、各重要行业正常运转不可或缺的基础设施。关键信息基础设施是行业运转体系中被强依赖的关键节点，它所承载的业务对其他部门或行业核心业务有较大关联性影响。对这类关键信息基础设施的攻击所产生的破坏，通过关联的行业、领域逐渐传递，会造成连锁连片的严重后果。因此，社会的持续运转需要大力加强对关键信息基础设施的保护。

三是对关键信息基础设施进行法律保护是顺应国际形势的必要举措。目前各个国家均已建立关键基础设施保护的相关制度，美、德、英、日等国家通过出台和发布政策、法律、标准等多种措施，构建了国家关键信息基础设施保护体系。各国通过发布或升级监管框架、出台指南、完善机构设置等方式进一步推动关键信息基础设施的安全防护工作落地和具体化，提升工业信息安全防护水平。而针对新一代信息技术的应用可能面临的信息安全风险，也有一些国家已做出了相关的尝试，如英国政府致力于保护关键基础设施免受针对计算机或通信系统的电子攻击威胁，并建立了由国务大臣负责的国家基础设施保护中心为核心，各基础设施部门具体实施相关职责的关键基础设施保护管理体系。因此，为了提升我国关键信息基础设施防护水平，加强监管，防止安全事件发生，我国须加快对关键信息基础设施相关法律法规细则进行研究制定工作。

四是加强关键信息基础设施保护对于公民福祉的保障意义重大。加强关键信息基础设施保护的根本是对公民福祉、公民利益的保护。关键信息基础设施运行过程中存储或传输的信息数据大量集中或极其敏感，其中供水、供电、医疗卫生、社会保障等公共服务领域的信息系统、政务网络及网络服务提供者所有和管理的网络及系统中有大量的公民身份信息、金融信息等，这些信息一旦被恶意收集或利用，必将损害公民的利益。基于其他行业对于关键信息基础设施的依赖性，加强关键信息基础设施的保护，可以使得公民的工作、生活等更加便利。国家安全、社会稳定及社会的持续运转等是公民福祉得以保障的前提，故加强关键信息基础设施建设也关乎公民福祉。

详细规定了运营者的责任和义务

《条例》在总则部分对运营者责任作了原则规定，要求运营者依照本条例和有关法律、行政法规的规定以及国家标准的强制性要求，在网络安全等级保护的基础上，采取技术保护措施和其他必要措施，应对网络安全事件，防范网络攻击和违法犯罪活动，保障关键信息基础设施安全稳定运行，维护数据的完整性、保密性和可用性。《条例》明确规定了运营者责任义务，主要包括：

一是建立健全网络安全保护制度和责任制，实行“一把手负责制”，明确运营者主要负责人负总责，保障人财物投入。

二是设置专门安全管理机构，履行安全保护职责，参与本单位与网络安全和信息化有关的决策，并对机构负责人和关键岗位人员进行安全背景审查。

三是对关键信息基础设施每年进行网络安全检测和风险评估，及时整改问题并按要求向保护工作部门报送情况。

四是关键信息基础设施发生重大网络安全事件或者发现重大网络安全威胁时，按规定向保护工作部门、公安机关报告。五是优先采购安全可信的网络产品和服务，并与提供者签订安全保密协议；可能影响国家安全的，应当按规定通过安全审查。与网络安全法相比，《条例》进一步明确了运营者责任义务，更具有可操作性。

便于运营者贯彻落实

总体来说，《条例》对运营者的要求更为明确和具体，便于运营者贯彻落实；同时运营者的安全责任也更加清晰，并确定到具体担责人，督促运营者主要负责人真正推动《条例》执行。根据《条例》要求，运营者需要重点做好以下三个方面工作：

一是落实主体责任，保障相关资金落实，建立网络安全保障体系。第一，建立安全保护制度，并保障资金投入。《条例》第十三条指出运营者应当建立健全网络安全保护制度和责任制，保障人力、财力、物力投入。第二，设置专门管理机构，壮大网络安全队伍，完善关键信息基础设施防护网络安全组织体系。运营者的主要负责人对关键信息基础设施安全保护负总责，领导关键信息基础设施安全保护和重大网络安全事件处置工作，组织研究解决重大网络安全问题。第十四条规定运营者应当设置专门安全管理机构，并对专门安全管理机构负责人和关键岗位人员进行安全背景审查。第三，加强网络安全意识教育。树立科学正确的网络安全观，常态化开展全员网络安全意识教育培训。

二是高度重视安全保护工作，合规做好系统建设相关工作。第一，保障相关资金落实，加强网络安全建设资金的集约化管理与使用，平衡分配网络安全建设投资。《条例》第十二条明确提出，安全保护措施应当与关键信息基础设施同步规划、同步建设、同步使用。第二，采购的产品和服务应当符合规定，并签订保密协议。第十九条规定运营者应当优先采购安全可信的网络产品和服务；采购网络产品和服务可能影响国家安全的，应当按照国家网络安全规定通过安全审查。第二十条规定运营者采购网络产品和服务，应当按照国家有关规定与网络产品和服务提供者签订安全保密协议，明确提供者的技术支持和安全保密义务与责任，并对义务与责任履行情况进行监督。

三是定期开展网络安全检测和风险评估，发生重大安全事件应及时报告。第一，应当定期开展网络安全风险评估。第十七条规定运营者应当自行或者委托网络安全服务机构对关键信息基础设施每年至少进行一次网络安全检测和风险评估，对发现的安全问题及时整改，并按照保护工作部门要求报送情况。第二，发生重大安全事件应当向有关部门报告。第十八条规定关键信息基础设施发生重大网络安全事件或者发现重大网络安全威胁时，运营者应当按照有关规定向保护工作部门、公安机关报告。