大数据安全能力实践

杜跃进，郑斌
阿里巴巴集团，浙江 杭州 310013

摘要
安全的目的是为了保障发展，如何衡量一个拥有数据的组织的数据安全保护能力是十分重要的。探讨了拥有数据的组织面临的数据安全问题及挑战，介绍了大数据环境下的数据安全发展趋势和完整的组织级数据安全能力框架，阐述了数据安全保护能力实现的路径及实践过程中可能遇到的难点。最后，以某互联网金融企业为例，分析了利用数据安全能力成熟度模型指导企业进行数据安全保护能力建设的过程和方法。

关键词
大数据；安全能力；成熟度模型；安全管理

中图分类号：TP309 文献标识码：A doi: 10.11959/j.issn.2096-0271.2017049

Security capability practice of big data

DU Yuejin, ZHENG Bin
Alibaba Group, Hangzhou 310013, China

Abstract
The purpose of security is to ensure development, and it is important to measure the data security capability of an organization. The security problems and challenges of data organization were discussed. The development trend of big data and data security capability framework were introduced. The path of data security capability implementation and difficulties in the process of practice were expounded. Take an internet financial enterprise as an example, the process and method of building data security capability by using data security capability maturity model were analyzed.

Key words
big data, security capability, maturity model, security management

1 引言

数据被称为新时代的“黄金”或者“石油”，正在成为企业的核心资产，成为创新的关键来源，成为国家的战略资源。数据越来越值钱，自然成为违法犯罪分子的重点关注目标。他们除了直接盗取数据进行倒卖之外，也会用全面的数据构建精准诈骗活动，甚至对用户数据进行加密，然后勒索赎金，这也成为了当今的主流攻击行为之一。在我国，以营利为目的的网络“黑灰”产业链活动从2004年底就开始了。随着网络中的应用日渐广泛和深入，犯罪分子能够攫取利益的地方也越来越多，因此团伙的人员规模也在不断膨胀。在网络“黑灰”产业链中，窃取用户数据是非常重要的一环。但是，直到2016年“徐玉玉事件”的发生才真正让我国全社会开始重视电信诈骗以及背后的数据泄露问题。随后，从各种不断披露的案例中可以发现一个现象：很多数据泄露都是通过买通内部人员来实施的，这完全不同于大家想象的“黑客范儿”。

2016年4月，欧洲议会通过了《一般数据保护条例》，并将在2018年5月25日生效。该条例对欧盟公民的隐私保护做出了极为严格的要求，违规企业可能最高被处以罚款2 000万欧元或者前一年全球总年营业额的4%。《一般数据保护条例》对全球众多企业都会产生非常大的影响。经过长时间的酝酿和讨论，2016年11月7日我国发布了《中华人民共和国网络安全法》，该法律于2017年6月1日实施。个人信息和重要数据的安全是这部法律的重要内容，相关的执行细则和标准（包括个人信息如何保护、数据跨境如何评估等）也在紧锣密鼓地制定。数据安全问题受到全世界从政府到普通消费者的各种不同角度的关注，但随着对数据安全的关注度越来越高，人们似乎正在陷入另外一种风险之中，那就是“数据恐慌”。这种“数据恐慌”表现为对数据采集和使用的过度限制或者禁止，而不是通过数据保护能力的提升来改善数据安全水平。如果这种趋势不能扼制，会导致法律法规、政策标准严重制约数字经济的发展，会使广大消费者对新经济丧失信心，从而导致各种创业创新严重受挫，这对于数字经济的发展是很危险的。安全的目的是为了保障发展，在目前的大数据应用和安全的环境下，非常迫切的一项工作是衡量一个拥有数据的组织的数据安全保护能力。

2 拥有数据的组织面临的挑战

数据只有流通共享，才能促进产业间协同，优化资源配置，更好地激活生产力。可以说，大数据时代下的生产过程就是数据采集、产生、应用、流通共享的过程，这是一个以数据为中心的经济时代，以数据为中心的安全能力至关重要。

大数据环境下，各组织机构都将面临着以下的数据问题及挑战。

（1）数据无处不在

伴随着信息化的开展，各组织机构的业务被大量数据化，数据被广泛应用于组织的业务支撑、经营分析与决策、新产品研发、外部合作，数据也不再只是管理者拥有的权利，上至管理者，下至一线业务岗位，都需要使用数据。

（2）系统、组织之间数据边界模糊

组织内部的核心业务系统、内部办公系统、外部协同系统不再是竖井式的架构，数据的共享使得各系统间存在大量的

数据接口，系统间呈网状结构，互为上下游，每个系统都是其他系统的一部分，同时，其他系统也是自身系统的一部分。数据的流通共享也进一步促进了组织间的协同，组织间的部分职能也互为上下游。

（3）数据关联、聚合更容易

大数据技术使得数据的采集、使用更加便利，数据的种类丰富，可被关联的数据要素大大增加，同时，运算能力的提升加大、加快了数据关联或聚合的效率和吞吐量。

（4）数据流动、处理更实时

实时数据处理技术的发展使得数据的流动和处理更加实时，在提升效率的同时，也加剧了安全的挑战。

（5）海量数据加密

组织内沉淀了大量的数据，涉敏数据量也远远超出以往的数量，传统的数据加密手段开始捉襟见肘，如何在灵活使用数据的同时，高效、安全地保护数据，也是需要解决的问题。

（6）数据的交换、交易

数据成为核心生产资料，其价值被高度重视，数据的交换、交易行为以及相关市场孕育而生，如何确保这些行为的安全，进而维护好国家、组织、个人的合法权益，是巨大的挑战。

（7）数据所有者和权利不停转换

目前行业里主流的数据相关方有数据主体、数据生产者、数据提供者、数据管理者、数据加工者、数据消费者，数据权利不停转换，而数据的所有者及相关权利的界定至今未能达成一致意见。

（8）业务的国际化

互联网化加剧了“地球村”的发展，网络虽然无国界，但是网络基础设施、网民、网络公司等实体都是有国籍的，各国虽然在网络主权的提法上各执己见，但在实践层面却无一例外对本国网络加以严厉管制，防止受到外部干涉。

3 数据安全能力框架

大数据环境下的数据安全具有五大趋势：从注重系统的防护到聚焦数据内容本身的保护、从单一组织的保障到跨组织的联动、从数据的保密到（大）数据经济秩序的保障、从技术风险+操作风险到技术风险+操作风险+商业风险+法律风险、从传统的数据技术到大数据技术。因此数据安全的能力必须充分考虑组织保障、管理政策及流程的落地、大数据治理、数据生命周期的安全、数据的风控、数据生态的安全协同六大要素。

如**图1**所示，数据安全能力成熟度模型（data security maturity model，DSMM）以数据生命周期为主线，聚焦数据安全相关的四大能力：组织建设、人员能力、制度流程、技术工具，对组织机构的数据安全能力进行评级，能够很好地帮助组织自身及合作伙伴评估数据安全能力，找到差距，有的放矢地提升数据安全能力，并作为数据共享的风险评判依据之一。能力成熟度等级维度组织的数据安全成熟度模型具有5个成熟度等级，分别是非正式执行（1级：随机、被动的安全过程）、计划跟踪（2级：主动、非正式的安全过程）、安全可控（3级：正式的规范的安全过程）、量化控制（4级：安全过程可控）、持续改进（5级：安全过程可调整）[①]。

4 实现路径与方法

（1）设立组织

为了有效保障数据安全政策的落地实

① http://www.cctime.com/html/2016-11-29/1246980.htm

施，企业应该设置专职的数据安全团队。此外，还需要设立面向全组织的数据安全委员会，委员会需要有来自业务、数据、安全、法律等领域的不同角色参与，形成专业上的互补和完整的组织视角，统筹全局的数据安全管理政策，兼顾发展与安全，推进各部门落实数据安全各项政策。数据安全是个系统工程，服务于组织的大数据战略，需要得到组织高层管理者的重视，数据安全委员会的负责人应该是组织里最高管理层里分管安全或者数据的管理者。

同时，还需要内部各相关部门的紧密配合。对于有多个业态的集团型组织，各业务的负责人应为该业态下数据安全第一责任人，与数据安全委员会、数据安全实体团队共同推动本业态下的数据安全工作。

（2）盘点现状

数据安全管理的核心是数据，需要对组织内的海量数据资产以及与数据相关的部门、业务/产品、流程、数据风险管理进行盘点。

数据资产的盘点：重点梳理数据的种类、数据量、核心的数据内容、数据来源以及数据的安全分级分类情况和流转链路。

数据相关部门的盘点：与数据相关的部门往往是数据风险的高发部门，属于高敏感岗位，需要梳理全组织与数据相关的部门数量、部门内部各岗位的职责、工作流程、数据操作环境，重点关注操作风险高的环节。

数据相关业务/产品的盘点：与数据相关的业务主要是指以数据为核心生产要素的业务，这类业务高度依赖数据，是组织对外提供数据服务的业务，在产品研发、测试和对外服务的过程中都需要对数据进行梳理，需要梳理数据在业务/产品中的应用原理、交互的系统接口、相关的责任人，此过程同样重点关注高风险的环节。同时，由于对外提供的是数据服务，提供的数据内容也需要进行合格性的盘点梳理。

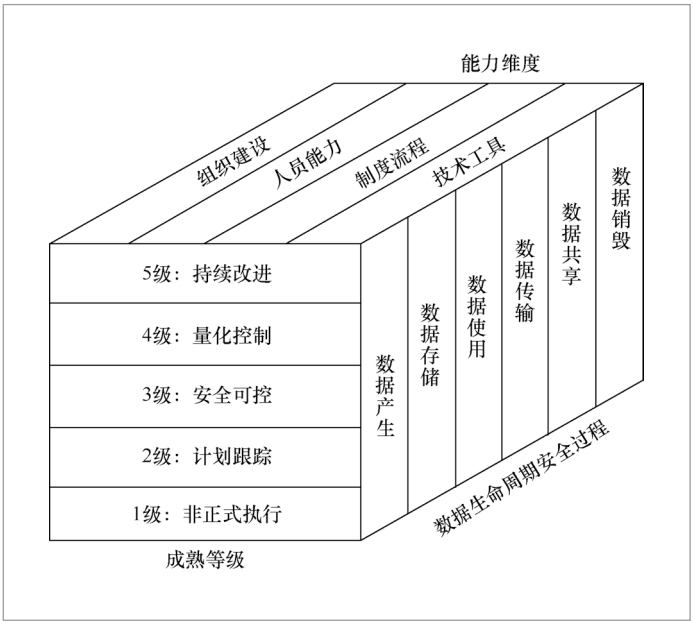

图1 数据安全能力成熟度模型

数据相关流程的盘点：数据相关流程指数据的采集、存储、授权、内部使用、传输、对外披露、销毁等过程，这些环节构成了数据在组织内部的主要流程，需要梳理所有线上线下的流程。

数据相关风险管理盘点：梳理数据风险的识别、风险评估及判定、风险跟踪及改进情况，包括基础性的治理，例如风险的日志数据、风险的定级机制、风险的响应机制。

（3）运用DSMM进行评估

如图2所示，DSMM包含32个安全域，涵盖组织的数据全生命周期过程，每个安全域含有相应的评估点和评估标准，由数据安全实体团队针对评估点参照评估标准进行安全能力评估。

（4）制定风险修复与短板提升计划

DSMM不但能够评估出数据安全能力，也能反映数据安全的风险，总体评估完成后，需要得到两部分的改进计划：一

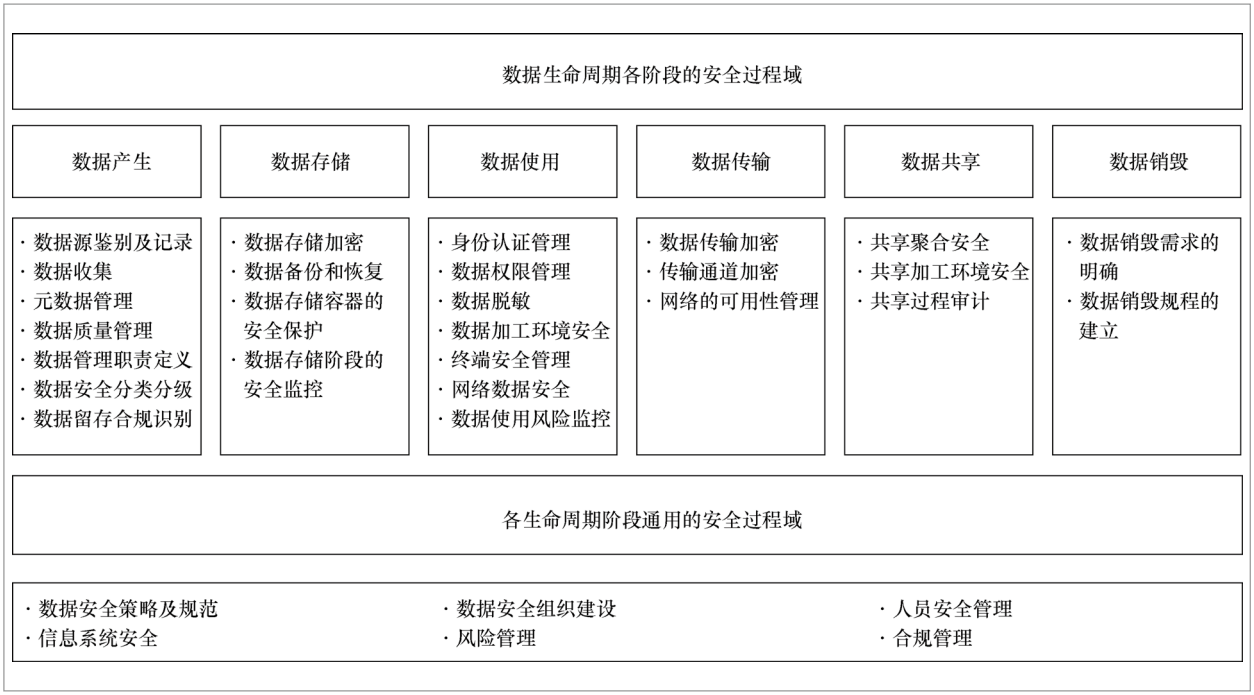

图 2 DSMM 的安全域

部分是风险修复计划，一部分是数据安全能力短板提升计划。

5 实践中的难点与挑战

在实践过程中，通常会遇到如下挑战。

（1）高层重视度不足

负责人的层级不够，难以协调；提供的资源投入有限，力度不够；仅仅作为合规需求，响应被动；缺乏前瞻性的布局，前瞻性的数据安全技术研究与投入缺乏或者不足。

（2）业务部门配合意愿度低

其他业务部门认为是安全部门的事情，主动性不强，业务要素的输入不足，导致数据安全政策不够贴近业务，既影响落地，又可能造成数据安全一刀切的局面，影响业务的发展。

（3）内部系统繁多，数据庞杂

业务的IT化促成了大量的系统产生，沉淀了大量的数据，应用系统的梳理、系统间的数据接口以及数据的盘点成为了基础治理工作的重中之重，日常实践中，基础治理工作往往得不到应有的重视，管理者往往急功近利，忽视基础治理工作的重要性。

（4）政策落地难

由于历史因素，组织里存在着大量的历史业务，大数据环境下的数据安全政策难免与现有业务流程产生冲突，冲突发生时的取舍容易导致数据安全为业务让路，造成数据安全政策落地难的局面。

（5）业务快速发展

“互联网+”或“大数据+”引发业务创新的加速，业务出现快速发展的势头，频繁迭代升级，数据安全政策及技术手段更新容易滞后。

（6）组织的关联公司多

大数据环境下，组织间的业务合作促进了数据的共享，如何安全可控地分享数据是大型组织常见的挑战。

6 案例分析：某互联网金融企业

6.1 企业概况

该企业融合“互联网+金融+汽车”，以互联网为主要渠道，为借款人与出借人实现直接借贷提供信息搜集、信息公布、资信评估、信息交互、借贷撮合等服务。车贷作为该企业的核心产品，其业务模式已经具备一套标准的流程，从自建工具实现贷款的线上操作管理，到自建车辆评估和全球定位系统（global positioning system，GPS）管理系统，实现数据化分析管理。在深耕车贷细分市场的同时，开启信用贷款、汽车消费金融、供应链金融等多个领域的持续性深度探索，逐步搭建以数据为核心生产资料的产品体系，有效提升了行业竞争力。

6.2 企业数据概况

主营业务中借款人与出借人的基本数据、车辆信息、与信用相关的数据、借还款行为数据、债权数据成为了业务的核心数据，数据概况见**表1**。

6.3 数据安全最紧迫的问题

该企业拥有几百万借款人和几十万投资人信息，近年来安全法律法规相继出台，监管日益趋严，满足监管及合规、保护个人

表1 某互联网金融企业的数据概况

数据类别	数据量	核心内容信息	数据来源
借款人	几百万	身份证、行驶证、机动车登记证及其影像 发动机号、车牌号码、行驶公里数等车辆信息及人车合影、车身照片、中控照片、里程表照片、发动机照片等影像 婚姻状况、学历、单位电话、月收入、工作单位、工作年限、单位地址、户籍所在地、现住址等个人信息 家庭成员及其他联系人联系信息 客户手机号、服务密码及近6个月运营商账单、详单（通话时间、通信地点、对方号码、通信时长等）、套餐、充值、短信信息等 客户公积金账号、密码及公积金系统个人信息、缴存信息、贷款信息、还款信	客户提供爬虫、系统生成