信息安全：大数据分析与数据治理思考

CPDA 数据分析师 [北京] 黄彩霞 （易道信安（北京）科技有限公司）

2017 年 6 月 1 日，我国网络安全领域的《中华人民共和国网络安全法正式实施，与此同时，网信办于 2017 年 7 月 11 日发布《关键信息基础设施安全保护条例(征求意见稿)》、全国信息安全标准化委员会组织于 2017 年 12 月 29 日正式发布国家标准 GB/T35273-2017《信息安全技术个人信息安全规范》并于 2018 年 5 月 1 日正式实施。

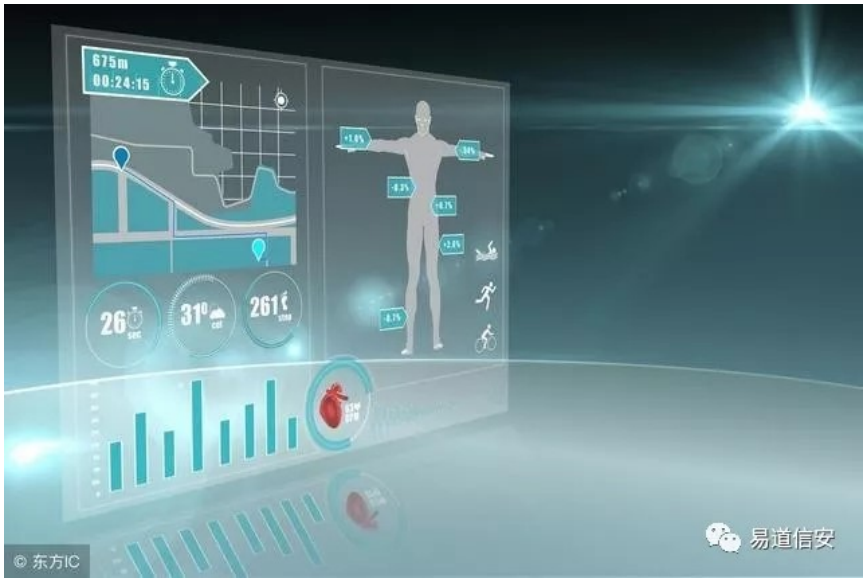
上述一系列法律法规和相关政策的颁布，标志我国网络安全、信息安全领域的法律监管迈入新的台阶。随着《网络安全法》的实施，它对网络安全，尤其是维护网络空间主权和国家安全、社会公共利益，保护公民、法人和其他组织的合法权益等方面，提出了新的要求。习近平总书记在网络安全和信息化 419 工作座谈会的讲话内容，对整个网络安全产业和相关领域的发展都有重要意义。

当前，大数据时代已经进入了人们的工作、生活、学习中，科技的高速发展给人们带来生活便利的同时，信息安全要想做好，它需要与大数据分析、数据治理之间的关系密不可分。

Sophos 首席数据科学家 Joshua Saxe 说：“我认为信息安全从业人员应该像数据科学家和黑客一样地思考。”因此，在这里，就信息安全领域，大数据分析与数据治理的逻辑关系进行简单思考。

一、多维化数据分析观

数据分析，不应该仅依附在传统的日志分析数据，应站在人性本身展开思考，通过对人的基础信息，赋予不同的权重，建立算法模型。不同的层面人所处的环境以及思维模式是不同的，所以在分析日志行为的时候，不应该局限于传统的算法模型，有过实战经验的数据分析从业人员都会有过深刻的体会，数据不等同于真实，如何去分析鉴别数据的真假，如何在海量级的数据指数进行合理的升维、降维分析，这是今后每个数据分析师要考虑清楚的，也是最考验数据分析师水平的砝码。



二、数据分析与工具应结合安全业务本身进行数据治理

在我们建立一套数据保护方案的时候，不难发现实施过程中，是一个比较繁琐且生命周期较缓慢的过程。数据提取及处理，要经历一个清理脏数据的过程，通常采用 ETL 的方式，也就是说数据抽取 Extract、数据转换 Transform、数据加载 Load，一般要消耗半年的时间。大数据安全相关人员要想解决这个难题，需要对于目标主体进行科学的分析，不仅是要考虑快速访问数据时效问题，还要充分考虑到安全风险控制方面的政策及高层领导决策层的要求。对于数据方面的安全审计，要对安全业务和数据分析、数据治理、数据科学等进行深入研究，建立一套分析的工具与平台进行兼容，这不局限与传统的软件，甚至可以集成到硬件设备里，实现大数据安全的有效防御和治理。

三、数据建模与数据治理的商业价值

通常提到数据治理方面，一般企业相关的负责人都会反问一句：做数据治理的价值是什么？我们为什么要请分析师帮我治理？我们现在的数据很好，并且数据这块我们目前也没有那么大的需求，对于企业本身来讲，请专业的人员做数据分析成本比较高，而数据治理需要分析能力很强的人进行挖掘最终呈现会需要一段的很艰辛的过程，它不仅需要各个部门进行配合协作支撑解决信息孤岛问题，而且要充分考虑各数据之间的逻辑关系，最重要的是要和决策者能达成共识，对关键数据的指标进行一个很明确的判断，因为不同的人站在不同的角度考虑的问题是不一样的，关键数据的指标确认以及赋予的权重是截然不同的。种种原因，导致数据治理方面，从一个好的创新创意方案到真正实施落地，需要很长的路要走。缺乏相关的考量，原来的数据治理企业部门是冷门，但近几年随着网络安全法的落地，越来越多的企业在国家的倡导下，开始对数据治理引起了重视。在我看来企业数据治理，重点从几个点开展：一个是围绕业务本身进行思考，另一个是高层决策者关注的核心指标。

（1）企业业务本身的思考：“业务”通俗的讲，是各行业中需要处理的事务，但通常偏向指销售的事务，因为任何公司单位最终仍然是以销售产品、销售服务、销售技术等等为主。“业务”最终的目的是“售出产品，换取利润”。就拿信息安全领域来讲，一般安全公司关注的业务本身的数据，比如研发的安全产品投放市场后给企业带来的利润。

（2）高层决策者关注的核心指标：这个需要建模数据分析师具备一定的沟通能力，通过多种途径去了解高层决策者所期望未来公司发展的趋势。像很多公司的决策者一般都关注的核心点会在开会的过程中，或者公司战略调整的时候会有相关的文件通知。要学会观察公司内部的人员安排以及公司的战略发展趋势，公司的时下热点以及对外宣传的消息。

大数据分析和数据治理是管理闭环的重要环节。信息安全分析时，不仅要站在全局的角度，从外部层面去考虑推动数据治理的外围环境的因素是什么，而且要充分考虑局部的角度，尤其是哪些企业内部遗留问题会成为企业发展的壁垒。

数据治理管理体系，如何让企业家重视，如何结合市场营销来促进数据治理。比如数据从管理层面展开优化，责任与企业参与的员工的绩效考核有关联，而不是单纯的数据分析，数据治理结合数据应用，数据治理架构与企业决策支持的体系建设相互结合，这样可以以数据为导火线，将数据治理联合企业的管理、考核、市场、决策支持，无论是对于企业本身来讲的价值是不言而喻的。数据治理做好不仅可以推动整个企业的文化而且能更科学合理的判断企业的利益最大化，超值的投入和支出比，数据驱动企业发展，更智能安全的促进整个数据安全，同时也要充分考虑信息安全的问题，大数据一旦泄漏会对企业不利，会导致整个战略陷入被动状态。我们可以创新的形式，将数据驱动带动大数据信息安全，通过信息安全的方式带动数据治理的整体价值。