

大数据分析应用的三个方向

(来源：中国电子政务网，2020-01-09)

国内大数据应用已快十年，然而，对大数据应用的深入思考却没跟上，以统计分析的模式来想象大数据的应用，定格在信息获取方向上会影响在其它领域的应用创新。本文作者原国家信息中心副主任胡小明，从“大数据应用”的角度带来一些全新启发！本文将从信息获取、网络服务、数据整合三个方向探讨大数据应用，拓宽应用视野。

一、信息获取的大数据应用

1.1 信息获取与知识发现

对大数据最容易想到的应用是信息获取，大数据应用经常被视为是统计分析模式的延伸，维克多·迈尔-舍恩伯格与肯尼斯·库克耶合著的《大数据时代》从信息获取、知识发现的视角提出了大数据应用带来的创新：一是使用全部数据替代抽样数据，二是允许数据的混杂型而非精确性，三是重视相关关系而非因果关系。大数据创新开辟了知识发现的新思路，促进了科学研究工作的发展。

1.2 政府决策的大数据应用

大数据在知识发现中呈现出来的优势使政府产生用大数据实现决策科学化的构想，决策是领导者的行为，但是人脑无法直接使用大数据决策，数据所包含的内容要浓缩经过理解之后才能成为人脑中的信息，与人脑其它信共同参与决策。数据挖掘、统计分析都是对数据内容进行浓缩处理，形成人脑易于理解的信息内容，数据包含的信息要通过人脑在决策中发挥作用。

1.3 政府大数据应用的局限性

大数据应用对政府决策的改进并不明显，否则不至于长期纂写不出政府大数据决策案例集。政府决策应用大数据困难主要有两点：

一是缺少适用的数据源，政府自己并没有多少大数据资源，适合当即需要决策问题的外部大数据也极难碰到；

二是很多重要信息如国际形势、重大事件、管理能力、社会文化很难数字化，政府仅靠数据决策会有很大的片面性；

1.4 大数据适合特定领域的决策

大数据主要来源特定的业务渠道，渠道的局限性使大数据获取的信息也有相应的局限性，因此大数据不适合大范围的决策，对政府的宏观决策帮助并不大，但是在微观应用中会有很多成功的应用，如案件侦破等。电子商务企业经常利用业务积累的数据分析用户需求，依据客户浏览内容推荐新产品与服务。

二、智能网络服务的大数据应用

2.1 智能网络服务也是大数据应用

大数据局限于信息获取应用会忽略在服务中的贡献。产生大数据的业务本身是更基础的大数据应用。谷歌、百度、阿里巴巴、腾讯、亚马逊等公司是大数据企业，电信运营商、银行等也是大数据企业，这些机构的大数据应用与信息获取应用不同，它们关心的是提供服务效率而不是知识获取，智能网络服务的大数据应用系统直接针对数据操作，不需要提取信息。

2.2 智能网络服务直接使用数据

大数据智能网络服务系统直接处理数据，为用户提供服务结果，这种业务由计算机流程自动处理数据，系统完全依据数据办事，没有

人脑参与就能达到极高的处理速度，确保处理结果一致性不受操作人影响。

2.3 智能服务的数据资源是动态数据流

信息获取的大数据应用是一次性运行，数据是静态的，一旦获取了信息即交由人脑处理，计算机的任务就结束了，智能网络服务则是连续的不停顿的业务，只要用户有需求系统就要响应，手机支付系统就需要不停工作以保证支付的及时性。电信运营商的服务亦不能停顿，其数据来自手机不停地向基站发出连接信号。连续的业务需要连续的数据源，智能网络服务处理的数据是在服务中实时产生的，它是连续的数据流。

2.4 云平台数据资源使服务智能化

智能网络服务需要来自用户的服务需求数据，还需要以前存储的数据资源，谷歌、百度需要收集网站数据以备用户查询。地理导航系统需要地图数据才能按照用户的实时位置计算导航路线。系统能够使用存放在云平台上的大量知识资源，高速的网络系统加上云提供的知识资源使网络服务如虎添翼，将普通网络服务升级为智能网络服务，智慧城市的网络服务暨是智能网络服务。

三、3.1 政府大数据概念的变通

政府部门的大数据概念与专家概念不同，专家认为常规信息技术无法处理的规模数据才算是大数据，因此政府没有什么大数据；政府认为各部门的数据汇集起来就是大数据，政府把大数据概念变通为数据整合的概念，使政府有大数据工作可做，但是《大数据时代》提出的大数据理念已不适用这种变通的应用，要用数据整合的思维方式推动政府变通的大数据应用。

3.2 政府业务数据管理与服务的特点

政府日常工作最重要的数据是各部门的业务管理数据，部门需要收集与积累本部门业务的管理与执行记录的数据。政府业务数据处理是对当事人或事项的精准处理，是下一步管理操作的依据，业务数据是不可替代的。

3.3 数据整合提升政府精细化管理能力

未来十年电子政务工作重点是实现政府业务的精细化管理，数据整合是实现精细化管理的重要手段。政府的信息管理能力取决于数据的完整性与现场调用能力，来不及调用的数据等于没有数据，现场管理人员与用户的信息不对称会影响服务效率并增加受骗机会。数据整合类似数据库建设，实现各部门数据在语义上统一，优化数据的组织，提高相关数据的调用速度，实现数据对现场工作人员的及时提供，发挥数据的整体优势。

3.4 数据整合提高政府公共服务效率

各地政府都提出让公众“只跑一次”和“一网通办”的目标，其目的是节约用户时间，跨部门数据的调用速度是提高服务效率的关键，数据整合可以提高跨部门数据调用的流畅性，提高公共服务的最终效率。

3.5 数据整合与信息共享不能混为一谈

将数据整合与信息共享区别对待是提高政府数据使用效率重要环节，数据整合的目标是提高政府业务操作的效率，目标很明确容易见效。数据整合工作宜采取应用导向，急用先做，提高效益。

信息共享目的是获取信息来支持决策与研究，常常需要调用整个数据集进行数据挖掘，信息共享需要对数据集整体的使用，数据整合

调用是对特定数据的精准调用，两种应用区别很大，不宜共用一个平台。

四、对政府大数据中心的建议

4.1 大数据中心可持续的关键是效益

决策者要认真考虑大数据中心持续经营的问题，否则上马容易下马难，可持续生存的关键是经济上合理，总效益大于总成本才能生存，大数据中心普遍的问题是想做的事情很多，但具体效益目标不清晰，很多项目的效益设想是建立在随大流的假定之上，盲目性很大，失败率很高，寻求效益可靠的服务项目是大数据中心生存的关键。

4.2 以改进政府精细化管理为中心

政府管理与服务的精细化是未来十年的中心任务，关键是做好政府业务数据整合，提高基层工作人员的现场工作效率，提升公众的满意度和政府部门的满意度。大数据中心可在支持政府领导决策方面努力，但重点是提高基层工作效率而非决策分析，基层工作改进容易见效而决策分析却很难让领导满意。

4.3 不要盲目地收集数据

太多数据会成为包袱，耗费设施资源还耗费管理精力。没有应用方向的数据不必忙于收集，更别指望会有人来信息共享，互联网时代数据资源早已过剩，有效益的应用目标才是数据应用的稀缺资源。

4.4 从最有效益的数据整合应用入手

政府数据整合对提高基层服务效率更能发挥作用，信息技术改进操作的效果明显高于改进决策。数据整合可以按数据合作圈分步推进，利用率高的数据整合先做，利用率低的后做甚至不做，不必齐步走，好钢用到刀刃上才是好方案。

4.5 开放的大数据中心更有生命力

大数据中心要促进政府数据向社会开放，数据开放是增加公众获得感的重要内容，公众支持率高能够直接提升大数据中心存在的价值。开放的大数据中心可以与企业合作，订购企业的可视化产品支持政府工作，将大数据中心做成政府与社会共享的可视化数据展示平台会很受欢迎，社会影响力越大，大数据中心生命力越强。

原文链接：

<http://www.e-gov.org.cn/article-170361.html>，转载请注明。