

大数据环境下未来档案馆建设的新构想

张馨月

(中海油天津液化天然气有限责任公司 天津 300452)

[摘要] 在档案馆信息化建设过程中,大数据技术的应用实现了档案管理质的飞跃。文章分析了大数据环境下未来档案馆的基本特征分析,包括高速度、大容量、多样性、智能化以及数据挖掘特征,并在此基础上提出未来档案馆建设的新构想:一是资源互通共享,分级有效管理;二是提供多样化服务功能,优化服务质量;三是建立信息数据档案库,实现主动存储。

[关键词] 大数据;未来档案馆;特征;新构想

引言

作为数据科学时代的代表性事物之一,大数据技术是依托现代信息技术形成和发展的,并逐渐在人类社会生活的各个领域中得到广泛应用。大数据技术的应用则为未来档案馆建设提供新的发展思路,以下将重点探讨大数据环境下未来档案馆建设的新构想。

一、大数据环境下未来档案馆的基本特征分析

大数据在数据科学理论的指导下,改变创新模式和理念,发展大数据技术,深化大数据应用和实践,而行业大数据将是大数据最大、最佳的应用领域。

(一) 高速度

高速度主要涉及数据处理及挖掘速度。相较于传统数据技术,大数据技术具备一项特别显著的优势,即数据处理速度极快。在数据处理方面有一条著名定律——“1秒定律”,即数据处理结果要在1秒内得到,一旦超出1秒,那么数据的价值也就不复存在。未来档案馆面对的信息数据必然是海量的,因此必然要大幅度提升数据挖掘、查询、检索及处理速度。

(二) 大容量

近几年来,云计算技术、云存储技术等被越来越多地应用到档案管理事业中,巨大的数据信息量将会给未来档案馆带来严峻挑战。而大数据技术的应用能够较好解决这一问题,不但能对大容量数据进行有效处理,还能实现未来档案馆数据容量空间的无限扩充,真正实现档案资源文件化。

(三) 多样性

就海量信息数据而言,大数据技术既是一项最有效地处理工具,又能对多样化数据格式及数据来源进行很好地应对。从类型上来看,数据既包括关系数据,同时也包括非关系数据,因此复杂程度较高。后者主要是指来源多样的半结构化数据格式与非结构化数据格式,如电子邮件、文件、搜索引擎、社会网络等产生的数据。相较于目前的档案文件,未来档案馆将会存储更多各式类型的源文件,因而数据格式及来源的多样性亦是未来档案馆的一项重要特征。

(四) 智能化

对于大数据技术自身来说,其本身涵括的技术手段是多种多样的,例如可扩展云计算技术、大规模并行处理技术、电网分布式数据库技术等等,而由这些先进技术所形成的数据生态系统可以说是巨大而复杂的。此外,未来档案馆能够依托更发达的科学技术进行专家决策系统的构建,便于准确预测人类社会的发展趋势,因此智能化是未来档案馆最明显也是最本质的特征。

(五) 数据挖掘特征

从档案事业的发展趋势可以明确,未来档案资料的来源渠道必然越来越广。而档案在形成阶段会有海量信息出现,在未来档案馆中运用大数据挖掘技术是必然趋势。而具有数据挖掘特征的未来自来档案馆,其发展和想象空间也是非常巨大的。

二、大数据环境下未来档案馆建设的新构想

作为一个档案大国,几千年档案工作的历史沉淀在我国形成了丰富多彩的档案文化,也形成了一套独特的档案思维模式。大数据时代,档案工作不再是装订、缝补,档案利用工作也不再是你来我找。如果采取无所作为、固守原状的政策,那么档案馆将会失去未来的发展机会,甚至会失去存在的意义。

因此,档案人员应从数据的视角看待问题,以大数据理念推动档案馆的管理与服务。做到以下几点:

(一) 资源互通共享,分级有效管理

首先,未来档案馆将实现资源的互通共享。实现档案信息资源的互通与共享是发挥档案馆功能的关键性条件。受技术水平限制,目前档案馆还未完全实现这一目标。通过搭建互联网平台将各个档案馆联系起来,构建一个巨大的档案数据信息传输系统,有助于进一步提高档案馆信息化水平,加速信息共享。而要实现这一目标,前提是要对大数据技术展开更深入的研究与应用。因此,当前应当基于规范要求统一布置档案馆,对档案资料进行科学分类及管理,对于档案馆内部的电子文档也要进行规范,方便对其进行查询、使用及传输。

(二) 提供多样化服务功能,优化服务质量

未来档案馆将为广大用户提供更为多样化的服务功能,服务质量将会大幅提升。通过建立专家决策系统和应用人工智能技术,再加上未来档案馆本身具备的海量数据资源,因此未来档案馆能够不但能满足不同用户的多样化需求,更能基于社会和国家要求导向开展一系列预测、分析及决策工作。通过自动化筛选社会信息、自动生成档案资料,能够实现真正意义上的档案资源化,有助于档案馆更好地提供各项智能化服务。

(三) 建立信息数据档案库,实现主动存储

就信息档案而言,从概念上可以界定为通过国家认定的、依靠现代化科技手段筛选存储的具有保留意义的重要信息;而数据档案指的是基于现代化信息手段对不同时期以及重要节点的信息档案进行追踪分析,在此基础上保留下来的对社会、对国家有意义的信息数据。基于此,从信息档案与数据档案这两个概念入手,延伸出信息档案库与数据档案库的建设,未来档案馆要将传统的档案管理模式保留下来,但更重要的是要进行信息档案库与数据档案库的建设。

未来档案馆通过建立信息档案库与数据档案库,能够对过去传统的档案存储模式加以突破,实现信息数据由被动接收向主动收集、主从存储的转变,在无限度扩大档案存储规模的同时实现档案管理的有序性。不仅如此,通过分析档案的历史记忆特点可以发现,如果将档案资料视作“证据”,那么“证”就是指资源化的信息档案,“据”则是指资源化的数据档案。相较于传统档案资料,资源化的信息档案与数据档案将会强化档案的历史记忆这一特质,使其更具可靠性与真实性。总的来说,大数据技术在档案馆建设事业中的应用将会推动我国档案事业迈向更高层次,具有无限的发展前景。

结束语

大数据在档案信息化建设中的具体应用,使未来档案馆实现质的飞跃成为可能,它不仅使档案存储资源化,同时使档案从被动存储向主动存储迈进,达到互联网互通、资源共享,加快档案馆智能化建设的步伐等这一切必将推动档案历史文化传承、国家政务公开、深化社会服务乃至智慧城市建设和发挥其重要作用。

参考文献

- [1]何夏昀.大数据应用视域下的智慧档案馆建设[J].山西档案,2018(03):84-86.
- [2]杨文刚.大数据背景下档案馆的建设研究[J].山西档案,2016(04):67-68.