# 高校热点事件专题档案潜能挖掘

当前，为社会热点事件服务已成为高校档案管理机构的文化自觉与自信，也是高校档案宣传工作的新常态。近两年来，各地高校档案馆（室）紧扣国际国内热点事件，利用档案这一文化“母资源”服务重点工作，在纪念中国共产党成立95周年、中国人民抗日战争暨世界反法西斯战争胜利70周年、南京大屠杀死难者国家公祭日、红军长征胜利80周年、孙中山先生诞辰150周年等重大活动举行之际，不断推出精彩纷呈的热点事件专题档案，在国内外引起了强烈反响。高校档案界这种集中公布档案的形式，彰显了档案在重大纪念活动中的独特作用。

研发与利用现状

通过挖掘档案，一些时间上间隔很久、地理上分散的事件之间的联系可以被发掘出来，零散事件的类似之处也得以发现。近年来，高校档案管理机构通过找准历史档案与现实需求的结合点，大力挖掘档案信息数据利用潜能的案例可谓不胜枚举。

如在纪念中国人民抗日战争暨世界反法西斯战争胜利70周年之际，南京大学档案馆通过整理文献、挖掘资料，成功举办了“抗战中的南大记忆”档案图片展，挖掘选取的400件图片和史料，向世人展示了高校战线可歌可泣的抗战记忆；在2014年首个南京大屠杀死难者国家公祭日之际，东南大学档案馆举办的国家公祭日图片展，通过大量史料图片向公众揭示了1937年东南大学的前身——国立中央大学遭到日军轰炸时的历史场景，用档案回击和戳穿了日本右翼势力否认甚至美化侵略的谎言；2015年，同济大学档案馆挖掘馆藏近700幅档案图片，举办了“抗日战争中的同济大学”档案图片展，全面展示了同济大学自吴淞校园被日军炸毁到抗战胜利后返沪的办学过程；在2016年第三个南京大屠杀死难者国家公祭日来临之际，南京师范大学等举办“1937·德国汉堡友人与南京”图片史料展，展览展出了档案史料近200件，缅怀了约翰·拉贝等国际友人对南京人民的正义之举，彰显了铭记历史、热爱和平的主题。

大数据开启了智能时代，今后高校档案管理机构可以更准确地把握时代脉搏，运用更多更丰富的表现形式来赋予档案信息大数据以新的生命，在全社会的大合唱中唱响、唱好“档案声部”，做好服务大局这篇大文章。

难点分析

目前，高校热点事件专题档案大数据处理遇到的主要难点表现在以下两点：

1.如何利用档案信息大数据来发掘故事

海量档案信息大数据彼此之间充满着互动联系，如何表现这些关系，充分发掘其背后的真实历史，这取决于档案工作者的挖掘能力。档案工作者应当借鉴博客营销的写作技巧，努力将档案信息的问题热点化、功能故事化、形象情节化、发展演义化、专题系列化、篇幅精短化。档案工作者要学会运用大数据技术来挖掘档案信息中一个个生动感人的故事，让读者感到耳目一新、记忆深刻，进而使档案信息产生广泛的影响力。

2.档案自然语言的多样性、复杂性

文本挖掘与处理技术是大数据挖掘领域中的一个重要研究分支，其主要研究从非结构化、半结构化的海量文本信息中挖掘有价值的信息，处理其中可能隐藏的概念及其联系，并完成可能的知识发现。文本挖掘与处理的主要对象是自然语言，但自然语言的复杂性、多样性决定了计算机很难实现对其含义的完全理解。因此，探索大数据浅层语言分析，是目前可行的一条研究路线。例如，可先着力挖掘诸如新的特征词、有意义串、关键词等内容进行浅层语言分析，然后再对档案文本信息进行热点事件的抽取、分类、聚类、检索、话题跟踪等大数据处理。

新特征分析与线索依据

在现实生活中，用户需要的热点事件信息通常都是粒度较小、内容相关、题材新颖的信息。因此，在高校热点事件专题档案大数据库的构建中，必须紧扣追踪热点事件的“新特征语言”和“有意义串”等相关线索依据来建立大数据源模块。

含有所谓“新特征语言”的句子，一般具有“内容相关性”和“时序新颖性”的鲜明特点。“内容相关性”指句子的内容必须与主题密切相关；“时序新颖性”指句子传达了以前所没有涵盖的新信息。

所谓“有意义串”，是指在真实语言环境中灵活使用且包含具体语义的“字符串”。这种“字符串”的共性是：在语义上、统计上、结构上和语用上可分别满足一定的性质。由此可见，一个“有意义串”就是一个线索（一个有价值的社会现象）。这种线索往往包含着网民对当前社会各种现象的立场和观点，所以对“有意义串”的挖掘，不仅有利于进一步深化关键、频繁模式的分析，而且对建立更高层次的文本内容自动抽取、话题动态监测、机器实时翻译等智能集成应用机制，都有重要的促进作用。

综上所述，“新特征语言”和“有意义串”的发现，对高校热点事件专题档案大数据分析与追踪具有重要意义。其作为线索依据的主要作用有：一是可以有效地促进系统处理平台对中文自然语言的理解分析与特征提取；二是可以有效地改善和提升文本分类、聚类的集成机制；三是可以在大规模档案数据中自动分析出热点事件并提供事件话题的关键特征描述。

应用探索

首先，搜集、了解用户兴趣点能在很大程度上增强对热点事件的描述能力和查询能力，直接改善用户的服务体验，这对用户和档案管理平台都具有非常重要的意义。高校档案馆服务的受众以年轻大学生为主，大学生接受新生事物的能力极强。因此，高校档案工作者可在知悉该群体利用需求的基础上，抓住其兴趣点，在重要纪念日期间开展多种形式的热点事件专题档案活动。

其次，高校档案管理机构借助舆情分析工具，可着力提升高校热点事件分析的效率和科学性。利用舆情分析工具在高校档案大数据智能处理平台上建立分析模块，并通过开发可视化图形，将热点事件做成放射图，可让用户在放射图中查看自己所关心的某个项目的具体评价结果。如果用户关心的话题（事件）不在此图之列，还可通过关键词进行检索，即用户点击某个具体热点事件，可视化工具就会立刻将该热点事件的各项指标从各类大数据库中抽取出来，并形成形象的放射图。

第三，高校档案管理机构借助当下最火爆的AR（Augmented Reality，增强现实）技术，可将热点事件档案大数据源的挖掘与利用变得更加生动和富有趣味性。AR技术是一种实时地计算摄影机影像的位置及角度并加上相应图像、视频、3D模型的技术，其更像是大数据的一种视觉化呈现模式，为用户带来一个既有虚拟又有现实的“魔法世界”。在这种新型科技革命的支持下，数字信息和纸质档案两种信息相互补充与叠加，用户用手机“AR”一下纸质档案上对应的图片，即能呈现出动态立体的有趣画面，全方位升级用户欣赏热点事件专题档案的体验感。为了让高校师生获得最新热点事件发展的动态信息，高校档案管理结构应当尽快引入AR技术，将其作为吸引师生与档案馆（室）互动的核心业务。用户只需要打开手机APP扫描海报、明信片或身边的宣传栏上特定的图片或二维码，即可在手机上看到神奇的AR特效。档案馆（室）可以结合开发的产品本身，在手机上展现各种充满动感元素的视频场景。例如，高校档案馆（室）策划某个热点事件相关的展览时准备呈现某个“镇馆之宝”，首先就需要扫描该物品的3D模型，并对其特定位置进行关键帧标定，然后根据馆藏物品的背景信息分别准备对应的动态视频数据。这样，当用户使用手机APP扫描藏品，系统就会去寻找AR场景中与之最接近的关键帧，根据关键帧上的特征点，利用特定的算法找到AR场景中对应的视频，并将其展现到用户的手机上。

第四，高校档案管理机构借助VR （Virtual Reality，虚拟现实）技术，可开发休闲娱乐形式的档案热点事件发布平台。VR是指借助计算机及最新传感器技术创造的一种崭新的人机交互手段。例如，美国《纽约时报》的馆藏照片如今成为奥运虚拟现实产品中炙手可热的“明星”，其推出的全实景360度立体虚拟现实电影《现代奥运会》，通过VR设备可以让观众身临第一届奥运会的赛场，见证历史上奥运冠军夺冠的场景，使观众获得独一无二的视觉体验。高校档案管理机构可以借助日新月异的VR技术推出热点事件相关影片，吸引新生代们的眼球。高校档案管理机构在做好热点事件线上宣传的同时，也要用心做好线下产品，并着力将线上的用户和线下的档案馆（室）实体进行有机对接，有效扩大用户规模，加强与用户的互动联系。只要抓住了受众的心理需求，高校档案服务工作也能够产生轰动效应。

张倩

 中国档案2017-5-5