

多模态信息融合的知识服务

张鹤 黄倩

(武汉科学技术情报中心文献检索部, 湖北武汉 430023)

摘要: 本文概述了大数据环境下多模态信息融合的开发和利用, 揭示了大数据背景下多模态信息的内涵、多模态信息融合的定义及其实现知识服务新的增值点, 阐述了关于多模态信息融合向知识服务细化延伸的几点思考, 进一步剖析了基于多模态信息融合的知识服务在满足用户深层次信息需求、实现知识服务创新与推动信息服务机构进一步发展的意义。

关键词: 大数据; 多模态信息融合; 知识服务; 用户信息需求; 信息服务机构

中图分类号: G203

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2016.06.008

Knowledge Services Based on the Perspective of Multi-model Information Fusion

ZHANG He, HUANG Qian

(Department of Literature Retrieval Wuhan Information Center of Science and Technology, Wuhan 430023)

Abstract: This article summarizes the development and utilization of Multi-model Information Fusion in the big data environment, Reveals Multi-model Information connotation, the definition of Multi-model Information fusion and its implementation knowledge service new value point. under the background of the big data environment. Elaborated the extension of thinking on Multi-model Information fusion to refine knowledge service, further analysis of the knowledge service based on multimodal information fusion in deep information to meet user needs, achieving knowledge service innovation and promote the further development of the information service.

Keywords: Big data, multi-model information fusion, knowledge service, user information needs, information service agencies

1 引言

随着互联网海量信息资源的快速增长, 过去单一的文本信息形式已逐渐发展为文本、图像、视频、音频等多模态的信息形式, 数据信息总量的提高导致各类互联网服务也应运而生, 并且越来越多地服务于广大信息用户。人们已经不再满足于单一的文本信息的获取和使用, 取而代之的

是更加智能化、个性化、多样化的信息与知识产品服务, 如获取文本、语音、图像等多种信息形式综合的多模态信息融合产品服务^[1]。

进入 21 世纪, 信息网络技术的不断发展, 催生了云平台、云存储、物联网等技术, 互联网信息资源呈现出“多模态、多维度”的特点。从信息资源的类别来讲, 互联网的信息资源种类已经远远超越了传统的文本信息, 图像、视频、音

作者简介: 张鹤*(1989—), 女, 武汉科学技术情报中心文献检索部初级工程师, 研究方向: 科技发展与政策; 黄倩(1985—), 女, 武汉科学技术情报中心文献检索部中级工程师, 研究方向: 科技政策与管理。

收稿时间: 2016年8月31日。

频信息在互联网上大幅增长；从用户的信息需求转变来讲，信息用户由以前传统的关键词搜索逐渐演变为更加深层次的语义检索，决策支持类的信息产品比简单的检索结果堆积更能满足信息用户的需求；从用户角色转换的角度来讲，大数据背景下，信息来源于多个方面，互联网信息传播是“众对众”方式，用户不但是信息的接受者，也是信息的发布者^[2]。从信息用户检索方式来讲，用户已不再满足于过去单纯的文本输入形式，检索方式开始向图像、视频、音频等多模态的形式转变，互联网信息服务机构也逐渐衍生出更具多样性的知识服务形式来满足不同用户的信息需求。

Web 2.0 环境下新兴的互联网社会化媒体不断涌现，宽松的信息环境更加强调了用户的主体作用，鼓励互联网信息用户主动地参与到信息的发布与共享中来，使得大数据环境下互联网信息资源呈现出海量、碎片化的特征。大量的多维度、多模态的网络信息资源既丰富了信息用户的视野，也为多种信息资源的融合交叉处理带来可能性。本文拟对多模态融合的知识服务进行初步探讨。

2 多模态信息的定义和内涵

所谓多模态信息，是指对同一个描述目标的不同或者相同侧面的有关信息，这些信息在来源方式、表现形式、语义内涵、可信度、侧重点和存储结构等方面都不尽相同，围绕同一个描述目标进行不同层面和不同维度的相关表达^[3]。相比传统单一的文本表现形式，多模态信息在描述目标的维度和层次上具有更加明显的优势。随着用户信息需求的不断变化，不同信息用户对同一信息资源在需求层次和侧重点上也不尽相同，多维度、多模态的信息资源在满足不同用户的信息需求上具有更大潜力。

信息融合的理念在不同领域均有提出，生物世界中信息融合广泛存在，工程领域中的信息融合从本质上来讲是用数学和机器对生物体的信息功能进行相应的模仿和抽象。信息融合是生物体

所固有的一种本质特征，是生物体环境感知和行为行动的基础，其过程是生物体生存、进化和发展的基本能力要素^[4]。目前，信息融合是指一种信息处理的过程，利用按时序获得若干传感器的观测信息，通过计算机技术加以自动分析、优化组合来完成所需的决策和估计任务。从定义中可以看出，信息融合的信息来源是由各种传感器所得，而多源的信息正是信息融合的加工对象，信息融合的最终目标是实现决策支持和综合处理^[5]。

大数据环境下的多模态信息融合，从信息来源的角度来看，早期的信息融合主要针对的是不同信息源的各类信息，亦称为多源信息，强调的是信息来源的多样性。而大数据环境下的多模态信息是指对同一个描述目标不同或者相同侧面的有关信息。多模态信息突破了传统的时间和空间上的限制，是对不同信息源的多样化信息特征和多模态信息模式的涵盖与扩展。对同一信息源的不同解读也构成了多模态信息资源的不同维度与层次；从多模态信息融合针对的具体对象来看，早期的信息融合来源于不同传感器的信号采集，更多的是一种数据化信息，如传感器信号序列等。大数据背景下的多模态信息融合更多地针对在 Web 环境下若干信息对象实体^[6]，如不同的网页信息、数据库资源、社交媒体信息等。因此，从信息管理的角度出发，大数据环境下多模态信息融合是指利用计算机技术，对多模态信息资源进行多维度和多层次的挖掘与重新组合排序，目的是逐一清除信息内容中的不确定性，从而获取更具有针对性的多模态信息融合知识产品的活动。多模态信息融合既是一种信息处理过程，也是一种提供信息知识产品服务。

3 开展多模态信息融合的知识服务的意义

当今社会早已进入知识经济的时代，传统的信息服务方式和内容已经不能满足信息用户的需求。在大数据背景下，海量的数据资源使得用户在寻找所需的信息与知识产品类服务时面临更多的挑战，信息用户希望通过更加快捷的方式获取自己所需的信息。各类传统的信息服务机构也逐

渐向知识服务转型,多模态的信息资源极大地提高了知识服务的质量,使得用户可以深层次多维度地对信息资源进行使用,达到所需的要求,不断发展的网络技术也为基于信息融合的知识服务提供强有力的技术支撑。

(1) 满足用户深层次的信息需求

在多模态信息环境下的信息用户,因为地域、文化程度、知识水平等诸多方面的差异造成了用户的信息获取能力的差异,这就给信息组织者提出了更高要求^[10]。在过去传统的信息环境下,用户通过单一的关键词搜索寻找所需的信息与知识资源,资源类型也偏向于文献、文本等信息。在大数据环境下,数据的总量和更新速度都大大提升,描述资源的形式也是多种多样,海量的数据资源提高了用户信息与知识产品的获取难度,也降低了用户对信息和知识产品的获取效率。将多模态信息融合的理念引入知识服务中,其根本目的就是从根本上满足用户在海量的数据信息资源中准确定位所需内容的信息资源,真正实现知识服务与用户需求之间的无缝式对接,实现知识组合的有序化。首先,通过信息融合技术,可以对大量数据资源进行相应的处理,去杂取精,抑制无用的信息资源盲目增长;其次,通过多模态信息的融合,更好地优化知识组织结构,从而更好地为用户提供服务,满足不同用户、不同层次的需求。

(2) 实现大数据环境下知识服务能力的不断提升

大数据背景下知识服务能力的提升离不开知识的创新,互联网不同领域的与不同学科之间的交流合作,是促进网络信息与知识服务内容不断创新的重要途径。通过多模态信息融合的知识服务的提出,能够从知识的继承,知识的交流、知识的创新等方面促进网络知识资源的有序化流动。首先,从知识的继承来说,多模态的信息融合可以更好地映射出不同类型学术交流活动中产生的新的知识单元,通过一定的融合算法将其进行内容和特征的聚类,从而纳入已有的知识单元或形成新的知识单元,是知识服务内容创新的重

要途径。其次,从知识的交流来说,多模态的信息融合将信息资源的不同维度、不同层面进行相应的聚合,使网络信息与知识资源的传递具有多渠道、多方式的特性,在满足用户信息需求的过程中,也实现了知识的良性传播与扩散。最后,从知识的创新来说,知识进步到一定程度后,就很难再局限在单一的领域或学科中,跨学科、多平台的知识交流是实现知识创新的必然趋势,多模态信息融合的知识服务可将不同领域的知识进行重组、排序、融合,从而脱离固定的学科主体,实现知识的有效创新与应用。

(3) 推动信息服务机构的进一步发展

知识经济时代为信息服务机构的发展提供了广阔的空间,知识经济的理念使得信息用户对知识服务的需求不断提升,也对信息服务机构的发展提出了新的要求。在大数据时代,社会整体信息化水平的不断提高,信息资源内容的广泛性、信息资源形式的多模态化、信息用户的社会化都对信息服务机构面对社会发展和知识服务市场提出更高的要求。知识服务作为一种开放式、智能化的应用,结合多模态信息融合的方法与技术,能够在新形势下将网络上的信息与知识资源进行高效整合,为信息用户提供不同层次、更具针对性的服务。基于多模态信息融合的知识服务不仅扩展了信息服务机构知识服务的内容,而且创新了其服务模式与特点,对信息服务机构的进一步发展具有强有力的推动作用。

4 几点思考

知识服务属于一种智力型服务,侧重于创新性。知识服务的特点在于其支持一种较深层次的智力服务,可以有针对性地解决用户提出的问题,是信息服务发展到一定阶段的产物。知识服务的出现转变了以往单一的服务模式,更强调以用户为中心的个性化服务,将多模态信息融合的概念引入是实现知识服务新的增值点。多模态信息融合通过深入挖掘不同模态信息资源层次和维度特点,提供更具智能化的“打包式”信息产品和跟踪服务,既深化了信息用户对数据的利用层

次又有效地拓展了信息情报机构服务范围，从长久来看有利于信息服务机构知识类的产品创新和信息、知识服务效率的提高。

(1) 构建基于云计算平台的多模态信息融合服务

云计算思想的起源可以追溯到 1961 年 John McCarthy 在麻省理工学院 100 周年校庆上的一次演讲。他认为随着计算机技术的发展，未来计算能力也许能够像水电等公共资源一样流通和出售^[7]。目前，随着互联网技术的不断发展，分布式计算、网格计算、云平台服务等概念逐渐走入人们的视野。新兴技术的发展使得大数据环境下，海量数据计算与存储的能力大幅提升，网络环境下的资源共享与协同机制，大大降低了计算的成本，提高了信息与知识产品服务的使用效率。目前，云计算的概念还没有一个统一定义，一般来说，云计算是指通过计算机网络访问逻辑上的计算资源，而不是仅仅通过本地计算机访问的资源，数据由服务提供商存储在服务器农场，云提供商通过网络提供在线服务^[8]。由此可见，云计算的特点是在大数据背景下实现计算资源的动态部署与分配，可按需使用，即购即付。多模态信息融合在数据处理、数据整合与数据存储方面与云计算的思想不谋而合，因此利用云计算平台来实施多模态信息融合的服务应用具备一定的可能性。

在大数据环境下构建基于云计算平台的多模态信息融合服务，将海量的多模态信息资源依托云平台强大的计算能力，对多模态信息资源进行相应的采集与预处理。预处理包括对各种形式的多模态信息资源进行一致性的检验，归一化的处理，实现多模态信息资源基于特定目标的资源聚类。对多模态信息资源的特征提取可针对不同维度和不同层次的信息资源进行相应的重组和排序，根据用户的实际需求，将多模态信息资源与用户需求相匹配，选择适当的融合算法，向用户提供“打包式”的信息与知识产品类的服务。将多模态信息融合的机制融入到云平台的服务中，实现平台即服务，通

过云平台强大的计算与存储能力，向用户实时推送所需的知识产品类服务，支持用户决策，同时接受用户的使用反馈，便于信息服务机构在服务内容和形式上加以改进与创新。

(2) 提供基于多模态信息融合理念的用户需求预测型知识服务

传统信息服务是根据用户提出的委托需求进行相应文献资料的检索和提供，本质上是属于一种被动的服务方式，知识服务的理念则更强调一种主动性和预测性，信息服务机构可在分析用户需求特点的基础上，深层次挖掘用户的潜在信息需求，并主动将其转化为知识服务提供给所需用户，以此来提高自身服务的主动性和及时性。将多模态信息融合的理念引入到提供知识服务的过程中去，信息服务机构拥有大量的数据资源，在长期的工作中也积累了相当一部分信息用户，可将这些用户资源有效地整合起来，通过不同维度与层次的用户需求的描述和分析，构建不同信息用户的多模态信息需求数据库，针对不同的用户提供其所需的个性化的知识产品和服务。现有的数据挖掘与分析等技术的发展也为多模态信息需求预测型知识服务的产生提供了相应的技术支持。及时向用户提供深层次的预测型知识服务相当于充当用户的“外脑”，为其提供有效、科学、可行的决策支撑，进一步激发用户的信息意识。

在提供基于用户的需求预测型知识服务时可引入知识图谱的概念。不同信息情报机构针对自身用户特点建立一个多模态信息融合的用户资源库，通过深入分析用户信息需求的演化过程，将不同用户所需的信息或知识产品以图谱形式归纳总结出来，为下一步挖掘深层信息需求提供一定的理论支撑。对企业来讲，预测型知识服务可提高企业自身的竞争实力，及时发现企业发展规划中的优劣势，以便于在市场竞争中处于一个主动位置；对科研机构来讲，知识图谱的建立也有利于其更加清晰地掌握所研究领域的最新发展成果，不同领域间知识图谱的交叉部分可为科研机构提供更多的研究创新点。

(3) 建立针对地方优势产业的多模态信息融

合专题知识数据库

众所周知,不同地区的优势产业对整个地区经济的发展具有极大的推动作用,同时这些优势产业对技术创新与知识服务的需求相对于其他产业来讲也更多,信息服务机构应充分关注此类优势产业对行业内最新的知识产品需求,建立针对地方优势产业的多模态信息融合专题知识数据库。对企业而言,此类专题知识数据库有助于其快速把握行业最新动态,有利于企业发展思路的制定和技术产品方面的创新;对信息服务机构来说,专题知识数据库的建立也有利于为工作人员提供比较全面系统的多模态数据融合信息,减少今后信息检索与分析的工作量,进一步完善优势产业的相关信息资源也有利于当地政府统筹规划管理,促进区域特色产业的深化研究和发展。

例如武汉市目前围绕产业链部署创新链,突出信息技术、生命健康、智能制造三大产业,形成了一批具有国际影响力的产业集群、创新成果、国际标准和专利技术等,武汉市科技信息服务机构可针对这三大优势产业建立相应的多模态信息融合专题知识数据库,方便相关高新企业及时了解最新技术成果,帮助实现成果转化等问题。新型多模态信息融合数据库在数据来源的类型上比传统的数据库更加丰富,文本、图像、视频、音频及其综合处理后的数据资源均可成为多模态信息融合数据库的信息来源,同时多模态数据库信息资源的采集方式与要求也将比传统的数据库更加复杂。多模态信息融合数据库在检索过程中可运用多种模态相融合的信息检索方法,不再局限于单一的文字或图片检索,通过具有科学性与合理性的算法,将各种模态的信息资源进行相应的关联,通过初步处理打包成具有专题指向性的信息产品^[9]。用户在检索相关信息时,得到的不仅仅是片面的、单一的信息资源,而是具有不同类型、不同模态的专题知识网格,在满足用户当前信息需求的同时,也给予用户深层次信息需求的挖掘创造可能性。在建立此类专题知识数据库时要注重多模态信息资源的新颖性和全面性,及时更新各类数据库信息。多模态信息融合

专题知识数据库的建立与当地优势产业的发展可形成互为促进的良好循环,一方面优势产业的发展有利于多模态信息融合数据库资料的进一步完善,另一方面数据库的建立也有利于推动该优势产业知识服务工作的开展。

5 结语

随着时代的发展,大数据环境下知识服务的方式和内容也在不断进步,由传统的信息检索服务到个性化的知识与知识产品服务,互联网技术的发展和信息处理能力的提升使得海量数据与信息处理变得更加快捷和智能。本文主要从大数据环境下知识服务的角度出发,结合多模态信息融合技术探讨知识服务在服务内容和服务方式上的优化,将多模态信息融合的理念引入到知识服务中,既是用户的信息需求对知识服务类产品提出的新要求,也是知识经济时代知识创新与发展的必然选择。

参考文献

- [1] 王海峰,吴华,刘占一.面向互联网的信息处理[J].中国科学:信息科学,2013(12):1624-1640.
- [2] 王大玲,冯时,张一飞,等.社交媒体多模态、多层次资源推荐技术研究[J].智能系统学报,2014(3):265-275.
- [3] 段萱.基于多模态信息融合的数字图书馆信息集成服务研究[J].佳木斯教育学院学报,2012(9):390,397.
- [4] 韩崇昭,朱洪艳,段战胜.多源信息融合[M].2版.北京:清华大学出版社,2010.
- [5] 潘泉,王增福,梁彦,等.信息融合理论的基本方法与进展(II)[J].控制理论与应用,2012(10):1233-1244.
- [6] 曹高辉,王学东,夏谦,等.多模态Web信息融合系统构建及服务机制研究[J].情报科学,2014(9):8-13.
- [7] 张建勋,古志民,郑超.云计算研究进展综述[J].计算机应用研究,2010(2):429-433.
- [8] 夏谦.泛在环境下多模态信息融合机制与应用体系研究[D].武汉:华中师范大学,2013.
- [9] 宣守盼.多模态情感数据库的研究与建立[D].上海:华东理工大学,2013.
- [10] 田明明.基于多模态信息融合的知识空间构建研究[D].武汉:华中师范大学,2014.