

【编者按】科学性、创新性、导向性、预警性作为本刊的办刊宗旨,使我们始终关注国内外科技管理理论与实践研究的前沿动态,为促进我国科技政策与科技管理决策科学化水平的提高,提供一个开放的学术研究平台。因此,本刊也会不定期刊发一些新的观点、新的理念以及新的研究方法,与广大的作者和读者进行探讨与争鸣。本期刊登的《大数据环境下技术创新管理方法研究》是作者就大数据背景下,对信息技术与技术创新管理结合这种新理念和新方法进行的初步探讨,所构建的双向技术创新管理模型也是对现有技术创新管理方法的一种创新。

大数据环境下技术创新管理方法研究

朱东华,张 巍,汪雪锋,李 兵,黄 颖,马 晶,许幸荣,杨 超,朱福进
(北京理工大学 管理与经济学院,北京 100081)

摘要 随着大数据时代的来临,大数据问题迅速从技术层面上升到国家战略的最高层。以美国大数据研究与开发计划项目在美国国防领域的研究问题与应用入手,深入研究大数据环境下国家战略的新变化与新思路,思考并探索如何让我国技术创新管理工作迅速适应大数据环境,并有效利用大数据环境的机遇与挑战等问题。提出了大数据环境下面向技术创新管理的双向决策模型,整合传统目标驱动决策与大数据环境下数据驱动决策的理念及方法,构建了评估与预测和监测与预警的技术创新管理模型,以期提升我国技术创新管理研究在大数据环境下提取知识与观点的能力。

关键词 大数据;技术创新管理;数据驱动决策;数据挖掘

中图分类号 F273.1

文献标识码 A

文章编号 1002-0241(2013)04-0172-09

0 引言

提高自主创新能力,建设创新型国家是我国自十七大以来确定的国家未来发展战略的核心思想,而在胜利召开的党的十八大会议中,更是明确提出了实施创新驱动发展战略的伟大构想,并指出科技创新是提高社会生产力和综合国力的战略支撑,必须摆在国家发展全局的核心位置,以及以全球视野谋划和推动创新,提高原始创新、集成创新和引进消化吸收再创新能力,更加注重协同创新等重要方针政策。这是党中央国务院在深刻把握当代科学技术与经济发展特征的基础上,为我国未来发展打造核心驱动力所做出的重大战略部署。因此,在这一特定环境下,如何实现技术创新管理工作的创新驱动发展,是我国科技工作者亟待思考的重要问题之一。

另一方面,近年来,大数据引起各国政府、产

业界和科技界的高度关注^[1],2012年3月19日,美国总统奥巴马宣布启动大数据研究与开发计划(Big Data Research and Development Initiative),旨在提高从大型复杂数字数据中抽取知识与观点的能力,以帮助解决国家在科学与工程中最紧迫的诸多挑战问题,增强国家安全,实现教育与学习的转变^[2]。该项计划的提出将大数据研究上升到美国国家战略的最高层,以国家政府为核心的全球大数据战略竞争正式拉开序幕。

大数据时代的来临,标志着传统数据挖掘方法已经不再适应日新月异的数据环境,在数据采集、数据存储、数据分析以及可视化等诸多方面捉襟见肘。与此同时,各行各业对数据的依赖性有增无减,甚至以数据为基础的定量分析方法也有逐步取代耗时耗力的以专家为基础的定性分析方法的趋势。正

收稿日期 2013-02-20

第一作者简介 朱东华(1963)男,福建省三明市人,教授,博士生导师,研究方向:数据挖掘、技术创新管理。

是在这样的大背景下,传统的 技术创新管理 方法也面临着巨大的挑战,传统的数据挖掘方法与工具技术逐渐无法应付技术领域中迅速涌现的大规模数据,更无法实现对如此量级数据的实时处理与分析,从而丧失产生辅助决策依据的数据分析方案与能力。因此,从我国技术创新管理工作的实际需求出发,如何有效把握 大数据 环境下的大机遇与大挑战,应对 大数据 给传统技术创新管理工作所带来的从思维体系到技术方法的大变革,实现我国技术创新管理工作的决策创新,既是本文的立足点,也是本文期望解决的实际问题。为达到这一目标,一方面,本文将结合当前国内外的最新研究进展,并着重关注已公开的美国 大数据 研究与开发计划 项目研究内容^[2],深入剖析美国国防部与国防部高级研究计划局相关项目的研究目标与研究对象,寻求 大数据 环境下美国国防研究中的新思路、新理念。另一方面,本文以数据挖掘技术为核心,构建 目标驱动与 数据驱动 并行的双向技术创新管理方法体系,为我国技术创新管理的实际工作提供有力支撑。

1 研究背景

1.1 国内外研究现状

麦肯锡全球研究院将 大数据 定义为 无法在一定时间内使用传统数据库软件工具对其内容进行获取、管理和处理的数据集合^[3]。大数据 可以认为是 分析 的另一种表述,它是寻求从数据中萃取知识,并将其转化为商业优势的智能化活动^[4]。在大多数文献中,大数据 被总结为如下三个特征(即三个V):

(1) 容量大(Volume) 2012年,全球数据总量约为2.7~3.5ZB。而每秒钟通过网络传输的数据总量已经超过了20年之前互联网的总存储量。

(2) 产生速度快(Velocity) IBM将这一属性定义为数据进出的高速性,也可以将这一特性引申理解为对大数据迅速的处理能力以及决策能力的要求。

(3) 多样性(Variety) 大数据所包含的数据种类不仅仅包含常规意义上的关系数据库或数据仓库数据,还容纳了如图片、流媒体以及社会网络信息在内的全新数据类型。

此外,大数据的 信息价值大与冗余信息多(Value) 这一属性也被众多专家学者纳为大数据的第四个特征。

虽然,大数据 一词早在20世纪80年代就已经出现,但是一直以来,大数据 仅仅是作为数据挖掘技术定义中 海量数据 的另一种表述,局限于计算机学科中的 海量数据 挖掘与处理技术等研究,并未形成具有影响力及代表意义的突出研究成果或研究方案。2008年,《Nature》杂志出版专刊《Big Data》,基于多个学科的实际研究现状系统地介绍了 大数据 所蕴含的潜在价值与挑战^[5]。自此,大数据 正式登上科学研究的大舞台,并成为各个学科中的研究热点问题。而2011年,《Science》杂志出版的专刊《Dealing with Data》则标志着 大数据 时代的到来^[6]。

笔者于2013年2月12日通过Web of Science在SCI与SSCI数据库中以 Big Data 为检索词,在主题中进行检索,共获得232篇文献。其中,196篇的出版时间在2008年以后(包括2008年),117篇来自2012年。由此可见,大数据 研究在全球范围内的 新与热。本文通过VantagePoint文本挖掘软件生成了SCI与SSCI数据库中 大数据 相关文献的基于关键词的关联图谱(见图1)。

从图1不难发现,Big Data 与 Cloud Computing (云计算)是所有文献的最核心,而33个关键词中计算机与软件技术相关的关键词在一半以上,其中具有 大数据 文件管理与处理软件Hadoop与MapReduce更是出现在了关联图的显著位置。不过,我们还发现,Data Management (数据管理)、Data Grid (数据格)以及 Ontology (本体)三个关键词形成了聚类,出现在图的右下角(圈中所示),这可以认为是 大数据 管理相关研究的涌现。由图1可知,当前国外的 大数据 研究工作仍然集中在如何进行大数据存储、处理、分析以及管理的技术及软件应用上,而关于 大数据 与应用管理科学的结合,尤其是与技术管理科学结合的研究几乎为零。

本文还对这232篇文章作者所在的国家与其相关文章的关键词进行了关联分析,生成基于关键词的国家间关联图如图2所示。很显然,以美国为核

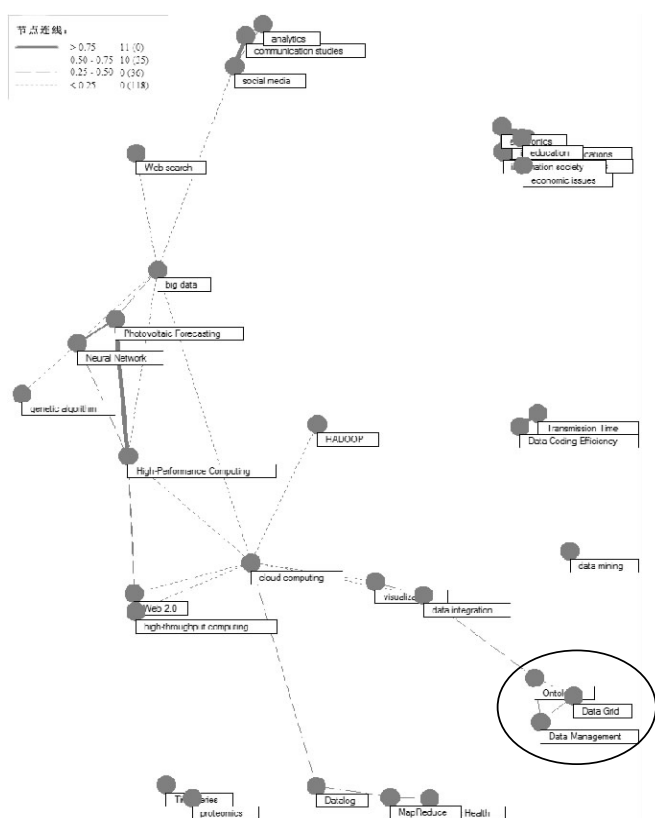


图1 SCI/SSCI数据库中 大数据 相关文献关键词关联图

心,包括德国、英国以及比利时等在内的欧美国家是研究 大数据 的最核心,而中国并不在这个核心集群之内,且与其他国家也几乎无关联。

放眼当前国内学者围绕 大数据 展开的相关研究,不难发现,虽然 大数据 研究已经引起了中国学者的高度重视,围绕 大数据 的研究工作也全面展开。然而,国内的 大数据 研究尚属起步阶段。自2012年5月份以 大数据科学与工程 一门新兴的交叉学科?为主题的香山会议顺利召开以来,在不到一年的时间,以李国杰院士为核心的计算机学科专家学者纷纷就围绕 大数据 定义的数据科学的理论架构与算法应用等问题进行了综述性的归纳研究与探讨^[1]。如马帅等就对 大数据 的异构数据模型与存储、复杂数据智能分析、数据质量以及大数据安全性等问题进行了分析和讨论^[7]。不过,针对如何应对 大数据 环境给科技管理问题带来的挑战,国内的相关研究微乎其微。在这一方面,中国科学院的王飞跃教授及其研究团队在处理与分析大数据使其有效支撑科技决策问题上进行了深入的研究,并

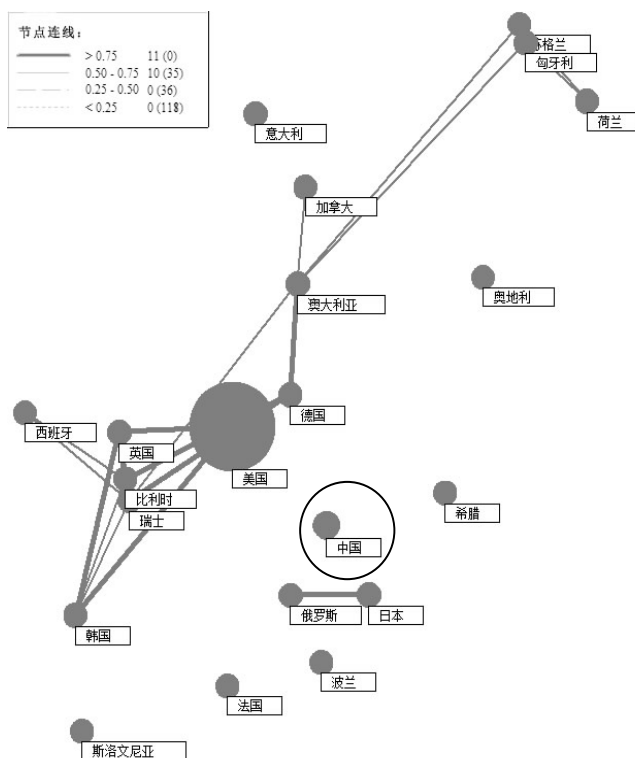


图2 SCI/SSCI数据库中 大数据 相关文献基于关键词的国家间关联图

构建了面向大数据和开源信息的科技态势解析与决策服务系统^[8]。

基于以上国内外研究现状,可以发现,大数据研究虽然是当前全球各个学科领域的研究热点问题,但是,相关研究仍然很大程度上集中在以云计算为核心的 大数据 相关计算机技术与软件的开发上,如何把 大数据研究 应用于实际的学科领域,或者说如何指导生产实践工作应对 大数据环境 的研究,几乎仅停留于理论阶段。因此,本文将着眼于美国的 大数据研究与开发计划,希望能从该计划的制定与实施过程中寻找能够将 大数据 研究与技术创新管理工作相结合的突破口。

1.2 美国 大数据研究与开发计划 与美国国防

自 大数据研究与开发计划 宣布启动以来,涉及该项计划的六大联邦部门(美国国家科学基金会、美国国家卫生研究院、美国能源部、美国国防部、美国国防部高级研究计划局与美国地质勘探局)承诺投资两亿多美元,以大力推动和改善与大数据相关的收集、组织和分析工具及技术,从而提升其从大型

复杂数字数据中抽取知识与观点的能力。其中,不难发现两个问题:一、国防部及其高级研究计划局均被列入大数据计划的核心部门之中,可见国家安全问题在本次美国大数据计划中有举足轻重的地位;二、从大型复杂数字数据中抽取知识与观点的能力是各大部门均强调的核心问题,而这也正是本次美国大数据计划尤为重视的预警与监测问题的根源所在。与此同时,美国国防部也提出应对大数据环境的基本策略为不断提高从数据到决策的能力,实现由数据优势向决策优势的转化。

美国国防部与国防部高级研究计划局针对大数据计划的立项情况以及各项目拟解决的核心问题(见表1),国防部每年通过2.5亿美元左右(其中6000万美元用于支持新的研究项目)的投资在各个军事部门开展一系列研究计划,旨在以创新方式使用海量数据,通过感知、认知和决策支持的组合,建立真正的、能够独立完成操控并做出决策的自治式系统,提高作战人员和分析人员的环境与状况感知能力,增强对任务和流程的支持^[9]。国防部的目标是将分析人员从任意语言文字资料中提取信息的能力提高100倍,同时希望他们观察到的目标、活动和事件的数量也获得相同幅度的提升。

本文针对表1所列美国国防部的8个项目和国防部高级研究计划局的XDATA项目进行分析与归纳,对其解决的问题目标之间的内在联系总结如图3所示。其中,XDATA项目是全部项目计划的基础,它通过提供基础技术与软件工具,实现对大数据的自动化处理能力;心灵之眼项目与机器读取项目则侧重于针对大数据的数据分析能力的提升,而

机器读取项目更将这种能力上升到了人工智能的高度;PROCEED项目集中于加密算法层,是大数据分析能力在具体领域中的应用研究;弹性云项目、ADAMS项目与CINDER项目则将大数据研究与国防紧密结合,把着眼点定位在网络之中,并通过相应的大数据监测能力,实现保障国家安全的最终目标;INSIGHT项目可以认为是国防部诸多项目的制高点,它将监测提升为预警能力,这也是数据驱动决策的最终体现。

因此,通过分析美国国防部针对大数据展开的项目研究内容及其目标,可以归纳为以下三个层次:面向大数据的基础工具与技术研发、以指挥与控制为导向的大数据分析及可视化、基于大数据研究的侦查与预警机制(如图4所示)。其具体内容包括如下几点:

面向大数据的基础工具和技术研发方面,基于非结构化与半结构化的大数据处理技术是美国国防项目研究的重点问题,以改善从国家快速增长的海量数据中提取知识与观点的能力。而军事数据中占据很大比例的军事图像及视频数据,又成为相关研究必须重视的问题之一。与此同时,信息安全问题也是国防研究中的重要挑战。

以指挥与控制为导向的大数据分析及可视化方面,美国希望运用先进的计算机技术极大程度地实现大数据指挥与控制的人工智能化,包括针对大数据的智能处理、辅助决策的能力。一方面,常规的计算机技术(如数据挖掘方法等)被大量的运用到针对大数据的智能化处理工作中来。另一方面,针对大数据的可视化研究也迅速成为众多研究的热点与

表1 美国国防部与国防部高级研究计划局 大数据 计划的立项及个项目拟解决问题

部门	项目名称	主要解决的问题
国防部	XDATA	开发处理半/非结构化数据的技术和工具
	心灵之眼	从流媒体数据中分析出作战信息
	机器读取	处理自然语言,实现人工智能
	PROCEED	针对加密数据的加密计算
	弹性云	监测、诊断和应对基于云计算的网络攻击
	ADAMS	海量网络数据的异常监测和鉴定
	CINDER	用创新方法检测军事网络间谍活动
	INSIGHT	提前确定网络威胁和不确定战争
国防部高级研究计划局	XDATA	开发处理半/非结构化数据的技术和工具

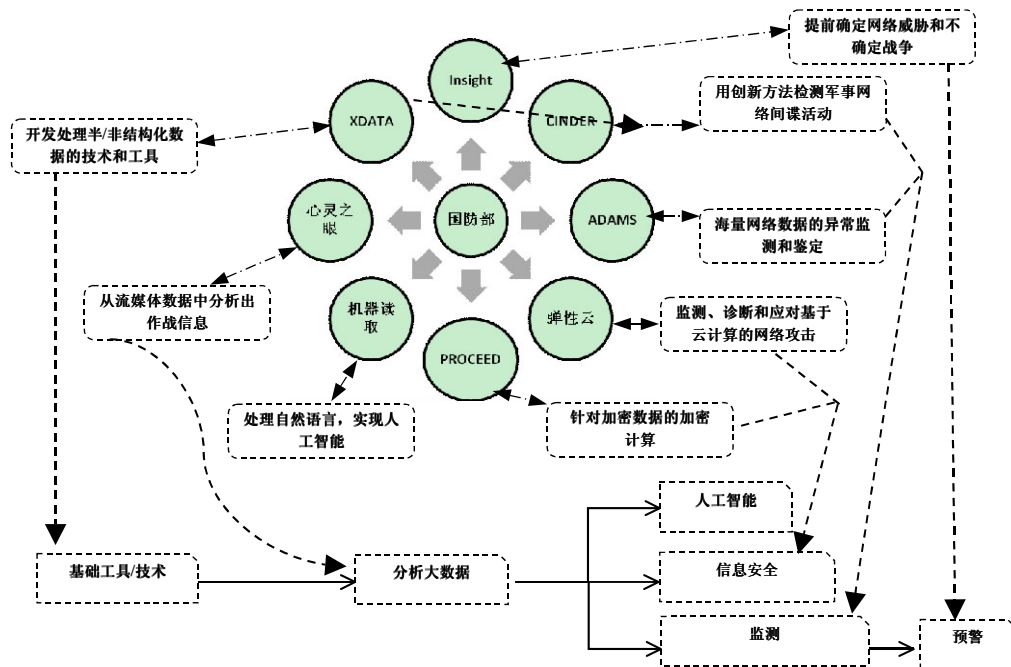


图3 美国国防部 大数据 项目目标内在关系图

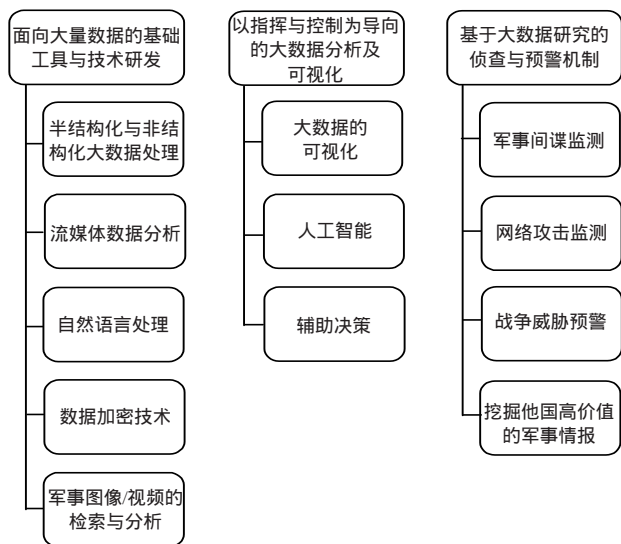


图4 大数据环境下美国国防研究目标层次图

重点。

基于大数据研究的侦查与预警机制 方面 ,提升军事情报的获取、跟踪、定位、处理、分析和预警能力 ,是各国军事研究中的最核心问题。美国为确保其在相关研究工作中的世界领先定位 ,更加关注于如何有效利用大数据这一挑战与契机 ,寻求通过有效的技术与工具提前识别军事情报中的潜在威胁 ,达到提前预警与防护的最终目的。

值得一提的是 ,在大数据项目的具体施行过程

中 ,美国国家政府有效地利用了美国科研机构与高校在基础算法与技术研究以及软件开发等领域的卓绝研究能力 ,与其展开深度合作 ,提升了项目研发的效率。例如 ,根据佐治亚理工大学官方网站上的相关报道 ,2012年11月 ,佐治亚理工大学工程学院的两个研究团队获得了一份来自于 大数据研究与开发计划 价值270万美元的技术研发项目 ,其侧重点为机器学习方法以及数据可视化方面的研究^[10]。

综上所述 ,基于美国 大数据研究与开发计划 项目的深入调研 ,尤其是针对大数据计划最核心的美国国防部及国防部高级研究计划局在相关立项中的侧重点与项目目标的分析 ,不难发现 ,从传统的 目标驱动决策 向 数据驱动决策 的重大转型是美国国家战略应对大数据环境的最直接变化之一 ,而基于传统数据分析方法的 评估与预测 理念也正在向 监测与预警 转化。有效地利用大数据环境 ,深入挖掘与分析潜在技术竞争情报 ,监测技术发展动态 ,分析潜在网络威胁与攻击 ,正是大数据环境下美国政府的全新国家战略思维与部署。

2 大数据环境下我国技术创新管理 双向 决策模型

如何应对 大数据 时代所带来的挑战与机遇 ?

美国 大数据研究与开发计划 项目提供了一条行之有效的思路与新方案,即转变决策驱动根源,由传统的 目标驱动决策 向 数据驱动决策 转化。以此为契机,我国的技术创新管理工作如何适应大数据环境,如何应对大数据环境下更为复杂与激烈的国际竞争,如何有效地将大数据环境作为重要契机运用到国家技术创新研究工作中来,正是本文希望达到的重要目标。

参照美国 大数据研究与开发计划 项目中引入的新思路,本文在整合技术创新管理现有理论方法的基础上,充分考虑大数据环境带来的挑战与机遇,构建了面向技术创新管理的双向决策模型,即目标

驱动模式与数据驱动模式并行的双向决策模型(见图5)。一方面,以传统目标驱动为基础,形成以传统数据挖掘技术与方法为核心的数据获取、处理、分析、支持决策的技术评估与预测模型。另一方面,围绕大数据环境下的复杂数据特征与环境,采用当前数据挖掘领域应用于大数据处理的新方法、新思路,形成以自组织动态实时监测为核心,能够有效预警并积极处理突发事件的技术监测与预警模型。

本模型的数据源均切实围绕 大数据 展开。一方面,包含科技文献(期刊、专利)、技术研究报告与统计数据等在内的科技数据是作为传统数据挖掘方法的核心数据源之一;另一方面,作为 大数据 环境

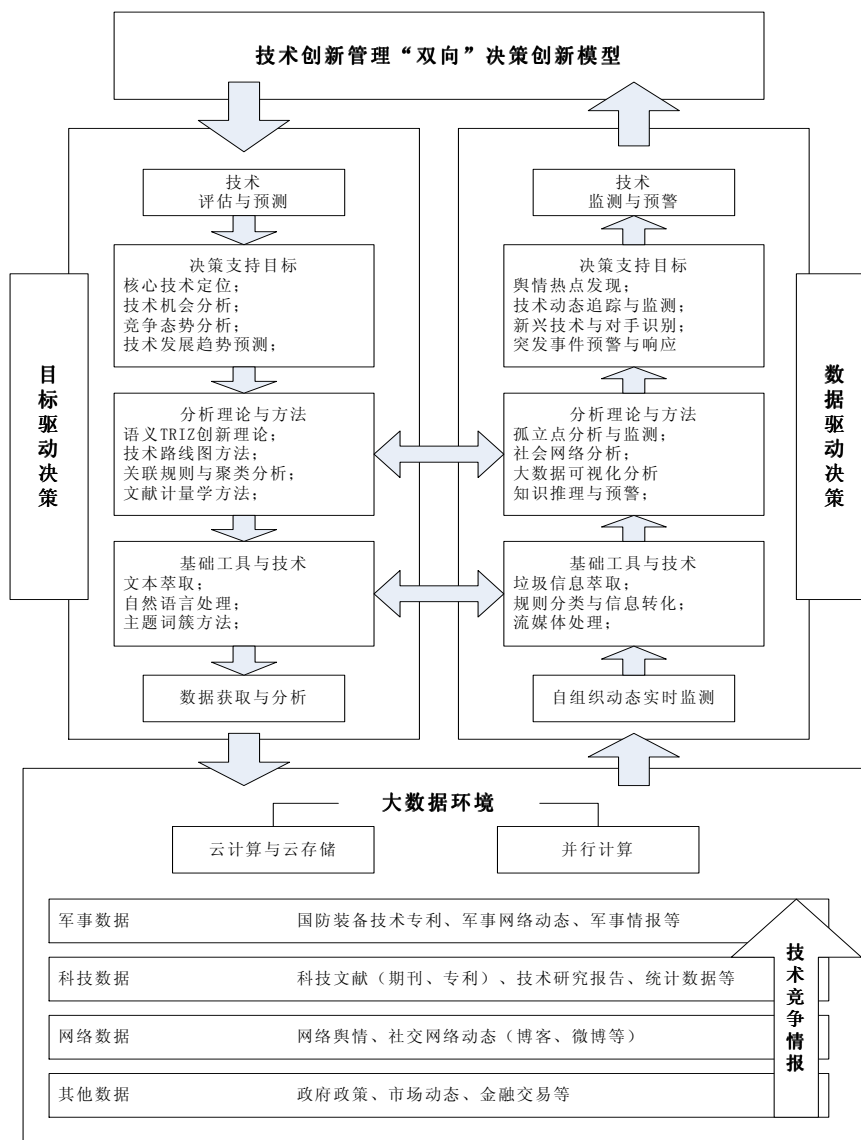


图5 大数据环境下技术创新管理 双向 决策模型

赖以生存的网络环境是考虑 大数据 数据源问题时不可忽视的问题,因此,本模型采纳了网络数据中的舆情数据与社交网络动态数据(博客、微博客等),这部分数据也是 大数据 环境下实现实时监测并预警的数据基础。此外,本模型还把政府政策数据、市场动态与企业数据以及金融交易数据等信息囊括在技术竞争情报分析的数据源中。而作为大数据环境中不可或缺的重要要素——云计算与云存储 技术与并行计算 技术也被本模型纳入到数据层中,作为本模型提取潜在信息、展开技术竞争情报分析的重要技术方法之一。

2.1 技术评估与预测模型

技术评估与预测模型围绕双向决策中的 目标驱动决策 展开,即决策者首先要明白自己的目的是什么,明白 干什么,然后,研究 怎么干。一般而言,目标驱动决策 面向具体的 技术 对象,在大数据环境下,采用常规的数据挖掘方法,通过数据采集、加工与分析计算等步骤,将数据转化为可供技术创新决策目标支持的有效知识或观点。具体运用在我国技术创新管理工作中,目标驱动决策 模型则通过有目的的对大数据环境下科技、网络以及其他数据中的潜在信息进行有效萃取与分析,实现对具体技术领域的 知识发现 与 可视化,从而达到评估技术发展状态,预测技术发展趋势的最终目的。

技术评估与预测模型围绕 技术挖掘 理论体系展开。其中,作为 技术挖掘 理论的核心思想

技术机会分析 方法由美国佐治亚理工大学 Alan Porter 教授于 1995 年提出^[11],即整合文献计量学方法与专家知识,针对与某特定新兴技术领域相关的科技文献数据展开挖掘与分析,定义技术领域的核心参与者、核心国家或地区、该技术的实时变化以及技术参与者或国家间的关联关系等。2003 年,Alan Porter 教授以 技术机会分析 方法为核心,提出 技术挖掘 理论^[12],系统地归纳了技术创新环境下如何掌握并分析科技文献数据^[13],提取 技术竞争情报,并通过 预测创新路径 方法进行可视化展示^[14]。本模型中采用了 技术挖掘 理论中归纳的技术

创新环境下进行技术竞争情报分析需要考虑的 39 个技术管理问题^[12],如 技术发展的热门方向是什么?、什么驱动了技术的发展?、技术的哪些方面符合应用目的?以及 哪些国家、公司、大学或研究机构在此技术或领域中领先?等。这些技术管理问题是结合实际技术创新管理应用,可供选择的实际 目标,而如何目标驱动决策,干什么 与 怎么干 由此展开。

技术评估与预测模型基于传统的数据挖掘方法,引入当前文献计量学与科学计量学领域中的适当新技术新方法,在兼顾传统数据挖掘能力的基础上,提升知识发现与表现能力。其中,基础工具与技术 部分在采用了传统的 自然语言处理 技术的基础上,引入了 主题词簇技术^[15],即综合运用停词表、模糊语义处理、主题词网络合并分析、词频与文本频次转换分析以及关联规则等文献计量学方法,实现对自然语言处理后仍然存在大量冗余主题词与噪声数据的数据清洗与合并工作。而 分析理论与方法 部分,本文则以 语义 TRIZ 理论 为核心^[16],打破传统词频分析专注于孤立的主题词的分析方法,通过语义分析方法,构建基于 主语 谓语 宾语的 问题与解决方案 模型,深入挖掘科技文献中的语义信息,并将知识发现结果通过技术路线图方法予以可视化表征。

2.2 技术监测与预警模型

大数据环境下,蕴含着丰富潜在信息的海量数据大量涌现,任何一种可能改变经济社会结构与现有技术水平的新兴技术或者任何一次可能产生技术创新的新思维、新想法都可能存在并以某种形式表征在大数据中。在国家技术创新管理工作中,如何掌控潜在新兴技术的发展动态,挖掘可能促进技术升级的核心技术,监测竞争对手相关技术领域的最新进展,并针对这一切可能情况做出及时并有效的反应,正是技术监测与预警模型所希望达到的最终目的。有别于技术评估与预测模型需要事先制定 目标,依 目标 的需求进行相应的数据挖掘工作,技术监测与预警模型更侧重于机器的自组织与自学

习能力,通过对小规模数据的训练与学习,形成反应并处理相应大数据的人工智能。这种自组织自学习的动态实时检测系统在传统数据挖掘方法的基础上,融入能够响应动态数据变化的动态数据挖掘理念,有效地运用相应的技术方法,能发现大数据中的容疑点,并深入挖掘与分析容疑点中的有效信息,从而达到最终形成监测与预警的最高目标。结合我国技术创新管理工作的实际需求,本文归纳了技术创新管理决策支持目标的四个方面,如表2所示。

表2 技术创新管理决策支持目标的具体问题列表

决策目标	具体问题
舆情热点发现	实时监测网络中的舆情热点,发现并追踪潜在技术创新思维与想法
技术动态监测与追踪	聚焦核心技术、动态监测技术发展问题,追踪热点技术、高精尖技术等
新兴技术与对手识别	发现新兴技术,定位新兴产业与新兴技术交叉热点,捕捉对手研发情报等
突发事件预警与响应	发现潜在技术研发策略或研发冲突等突发事件,及时反馈并产生响应方案等

有别于技术评估与预测模型中的传统数据挖掘技术,技术监测与预警模型中以数据为驱动,首先就表现在自组织动态实时监测模块的构建。通过规则的建立与基于机器学习的算法训练,形成专注于数据本身的实时监测模型。在剔除垃圾数据,实现数据向信息的有效转化之后,一方面,监测数据分析结果中的孤立点,及时发现异常数据并分析推演其背后原因;另一方面,基于有效的大数据可视化以及社会网络分析,聚焦并解决表2中决策目标的具体问题。值得一提的是,在技术监测与预警模型中,基础工具与技术与分析理论与方法两部分的部分理论方法是互通的,可以在实际的应用过程中,根据需求进行有效地选择与应用。

3 结论与展望

如同众多媒体及研究人员将2013年称为大数据元年一样,经过长期的积淀与迅速的发酵,大数据环境已经在短短一年的时间里扩展到经济社会的各个层面与领域,并从技术问题上升到了国家战略的最高层。如前所述,如何让我国技术创新管理研究工作迅速适应大数据环境,并有效利用大数据

环境产生的空前机遇与挑战,是每一名科技工作者不可避免并必须慎重考虑的重要问题。本文正是以此为突破口,鉴于相关研究团队在数据挖掘领域的多年积累以及在大数据研究中的深入思考,并充分结合在技术创新管理领域的丰富项目经历与研究经验,创造性地提出了大数据环境下面向技术创新管理的双向决策模型,整合传统的目标驱动决策与大数据环境下的数据驱动决策理念及方法,分别从技术评估与预测及技术监测与预警两个方面构建了适用于我国技术创新管理的新方案,旨在提升我国技术创新管理研究工作,应对大数据环境,有效并快速提取知识与观点的能力,并最终促进我国核心技术创新与升级,在激烈的国际竞争中立于不败之地。

与世界上所有国家、企业以及研究机构一道,针对大数据的研究,我们站在了同一起跑线上。如何抓住机遇,应对挑战,让我国科学技术的发展在这一迅速变化的复杂竞争环境中处于世界的最前沿,是本研究团队期望能与我国所有科学技术研究人员一同努力并达到的最终目标。在本文的基础上,进一步完善并丰富本技术创新管理决策模型,在开发智能化的基础软件工具的同时,选择特定的战略核心技术领域展开实证研究,是下一步需要解决的重要问题之一。

参考文献

- [1] 李国杰. 大数据研究:未来科技及经济社会发展的重大战略领域:大数据的研究现状与科学思考[J]. 中国科学院院刊,2012,27(6):647-657
- [2] Weiss R, Zgorski L. Obama Administration Unveils 'Big Data' Initiative: Announces \$200 Million In New R&D Investments[R/OL]. [2012-03-29]. Washington: Office of Science and Technology Policy, Executive Office of the President, White House, http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf
- [3] Manyika, J, Chui M, Brown J, et al. Big Data: The Next Frontier for Innovation, Competition and Productivity[R]. McKinsey Global Institute, 2011
- [4] McAfee A, Brynjolfsson E. Big data: The management

- revolution[J]. Boston: Harvard Business Review, 2012 (October):3-9
- [5] Specials Archive. Big data[DB/OL]. [2008-09-03]. Nature, <http://www.nature.com/news/specials/bigdata/index.html>
- [6] Special Online Collection. Dealing with data[DB/OL]. [2011-02-11]. Science, <http://www.sciencemag.org/site/special/data/>
- [7] 马帅,李建欣,胡春明. 大数据科学与工程挑战与思考[J]. 中国计算机学会通讯, 2012, 8(9):22-30
- [8] 王飞跃. 知识产生方式和科技决策支撑的重大变革 面向大数据和开源信息的科技态势解析与决策服务[J]. 中国科学院院刊, 2012, 27(5):527-537
- [9] 赛迪智库软件与信息服务研究所. 美国将发展大数据提升到战略层面[N/OL]. 中国电子报, 2012-07-17(3)
- [10] Robinson R. Georgia tech wins \$2.7 million to advance big-data technology for DARPA[EB/OL]. [2012-11-29]. <http://www.gatech.edu/research/news/georgia-tech-wins-27-million-advance-big-data-technology-darpa>
- [11] Porter A, Detampel M. Technology opportunity analysis[J]. Technological Forecasting and Social Change, 1995(49): 237-255
- [12] Porter A, Cunningham S. Tech Mining: Exploiting New technologies for Competitive Advantage[M]. New York: Wiley, 2005
- [13] Zhu D, Porter A. Automated extraction and visualization of information for technological intelligence and forecasting[J]. Technological Forecasting and Social Change, 2002(69):495-506
- [14] Robinson D R, Huang L, Guo Y, et al. Forecasting innovation pathways for new and emerging science & technologies[J]. Technological Forecasting and Social Change, to appear
- [15] Porter A, Zhang Y, Sakurai S. Text Clumping for Technical Intelligence, Theory and Applications for Advanced Text Mining[M]. Croatia: In Tech Publishing, 2012
- [16] Zhang Y, Zhou X, Porter A, et al. Triple helix innovation in China's dye-sensitized solar cell industry: Hybrid methods with semantic TRIZ and technology roadmapping [J]. Scientometrics, to appear
- (责任编辑 孙 兰)

Research on the Methodology of Technology Innovation Management with Big Data

ZHU Donghua, ZHANG Yi, WANG Xuefeng, LI Bing, HUANG Ying, MA Jing, XU Xingrong, YANG Chao, ZHU Fujin

(School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China)

Abstract: With the coming of the Big Data period, Big Data researches have escalated from technical level to national R&D level rapidly. This paper focused on the research questions and applications of 'Big Data Research and Development Initiative' program in American national defense fields, explored the new challenging and discoveries for national R&D strategy, and tried to apply these new ideas into China's technology innovation management researches in order to help handle with the opportunities from Big Data effectively. This paper proposed the two-way decision making model for the technology innovation management researches, involving both the traditional 'objective-based decision making' and the 'data-based decision making' concepts with Big Data, and extended this model to 'assessment and forecasting' and 'monitoring and early warning' systems separately. The aim of this paper is to improve the ability to extract knowledge and insights from large and complex collections of digital data in China's technology innovation management researches.

Key words: Big Data; technology innovation management; data-based decision making; data mining