

走向知识融合

——大数据环境下情报学的发展趋势*

李广建 罗立群

摘 要 本文首先对当前大数据环境下情报学研究与实践知识融合的现状和趋势进行了总结,分别从情报理念、情报采集、情报分析、情报服务四个方面梳理了情报学研究与实践有关知识融合应用的特点。其次概述现有知识科学领域中知识融合研究与实践的内容,并按主要研究内容和特征将其划分为三个发展阶段,分别为基于代理的知识融合、基于模式的知识融合、基于机器学习的知识融合。最后,结合近年来的情报实践,提出了一个知识融合研究的总体框架,一方面对现有情报学研究与实践进行概括和总结,反映当前情报学的新进展、新特征,另一方面勾画出未来情报学的一种发展路径,从而推动情报学的发展。图1。表1。参考文献58。

关键词 情报学 知识融合 情报分析 信息分析 大数据

分类号 G250.2

Towards Knowledge Fusion: The Development Trend of Information Science in Big Data Environment

LI Guangjian & LUO Lique

ABSTRACT

In recent years, an important development trend of the intelligence community is to emphasize the integration and consilience of different sources and types of intelligence. In this process, knowledge fusion plays a key role, which has attracted the attention of researchers and practitioners of information science and related disciplines.

This article first summarizes the current situation and trends of knowledge fusion in the research and practice of information science under the current big data environment: 1) The concept of intelligence has shifted from assisted decision support to direct prediction and early warning. In the process of warning and early warning, knowledge fusion is indispensable. 2) Intelligence gathering has shifted from traditional task-oriented passive gathering to active perception of content-based understanding. In the big data environment, with the development of data processing technology and intelligent algorithm, the method of information collection has transformed from traditional task-oriented collection to more intelligent intelligence perception. 3) Intelligence analysis has further expanded from focusing on association

* 本文系国家自然科学基金重大项目“大数据时代知识融合的体系架构、实现模式及实证研究”(编号:15ZDB129)的研究成果之一。(This paper is an outcome of the major project “The Architecture, Implementation Pattern and Empirical Research of Knowledge Fusion in the Big Data Era”(No. 15ZDB129) supported by National Social Science Foundation of China.)

通信作者:李广建,Email:ligj@pku.edu.cn,ORCID:0000-0002-2897-6246 (Correspondence should be addressed to LI Guangjian, Email:ligj@pku.edu.cn,ORCID:0000-0002-2897-6246)

relationships to the exploration of causality. In the big data environment, the demand for causal analysis in the intelligence field is increasingly strong, and it is gradually becoming the “normal” of intelligence analysis. 4) Intelligence services have changed from knowledge services to wisdom services. On the one hand, the application of intelligent information technology is the core of the smart service, which is different from the application of information technology in the past. On the other hand, smart intelligence services not only need to have a comprehensive and global perception of the intelligence environment but also have to deeply understand the intelligence environment issues by merging background knowledge and tacit knowledge of related issues.

Secondly, the article summarizes the existing knowledge fusion research and practice in the field of knowledge science, and divides it into three development stages according to the main research content and characteristics: agent-based knowledge fusion, pattern-based knowledge fusion, and machine learning-based knowledge fusion. Agent-based knowledge fusion is an early representative of knowledge fusion research. Its core task is to solve the problem of knowledge sharing, reuse and transformation in a distributed information environment, and to implement knowledge search and extraction of many knowledge resources through middleware technology. Pattern-based knowledge fusion mainly realizes the fusion of knowledge by changing and reorganizing the internal and external structure and attributes of knowledge driven by multi-source ontology. Its main goal is to solve the problem of contextual situation awareness and decision-making in complex scenarios. Machine learning-based knowledge fusion mainly uses automatic methods such as machine learning to achieve automatic knowledge extraction and learning of open data, the establishment of knowledge links, and the unification of knowledge. Finally, automatic construction of large-scale knowledge bases and automatic organization of knowledge are achieved.

Finally, combined with the intelligence practice in recent years, this paper proposes a general framework for knowledge fusion research, including thought domain, theory domain, technology domain, and application domain. The thought domain is our set of guiding thought for dealing with complex, changeable, and deeply uncertain intelligence environments. It is used to guide and lead intelligence research and intelligence work ideologically and theoretically. The thought domain of knowledge fusion in the field of intelligence is composed of fusion thinking, intelligence thinking, and computational thinking, and the three kinds of thinking complement each other. The main basic or source disciplines of these three kinds of thinking are cognitive science, information science, and computational science, respectively. The theory domain of knowledge fusion mainly includes four aspects: knowledge fusion theory, knowledge fusion framework, knowledge fusion model, and knowledge fusion method. Technology domain refers to different types of intelligence and cognitive activities. Knowledge fusion models, methods, frameworks, software and other technical elements are different. In this way, a set of technical elements oriented to a specific task or solving a type of problem forms different technologies field. The current application fields of knowledge fusion include financial intelligence, public safety, business analysis, military intelligence, scientific forecasting, and public opinion management. We believe that with the continuous development of technology, the application space of knowledge fusion will become wider and wider. 1 fig. 1 tab. 58 refs.

KEY WORDS

Information science. Knowledge fusion. Intelligence analysis. Information analysis. Big data.

0 引言

近年来,随着大数据和计算技术的发展,不论是情报理念,还是包括情报采集、情报组织、情报分析和情报服务等在内的各个情报工作环节,都在发生着革命性的变化,情报理论学者玛格丽特·麦克唐纳(Margaret MacDonald)认为,不断增加的信息量给情报机构带来的挑战是“情报界不再遭受信息短缺的困扰,而是遭受信息超载的困扰”^[1]。美国前国家安全局局长威廉·宾尼(William Binney)曾指出,大数据中的信噪比非常低,并且大量的数据收集使分析极为困难,甚至认为“国家安全局对大数据真是不知所措”^[2]。为应付这种新的挑战,近年来,情报界的一个重要发展趋势就是特别强调不同来源及不同类型情报的集成整合及融汇统一^[3]。在这个过程中,知识融合扮演着关键的角色,已经引起了情报学及其相关学科的研究与实践人员的重视。在理论上,情报界的许多学者对知识融合进行了深入的研究,包括联合情报^[4]、知识融合框架^[5]、知识融合框架模型^[6]、知识融合体系架构^[7]等内容。在实践上,美国国防部高级研究计划局(DARPA)和美国情报高级研究计划局(IARPA)等机构近年来开展了大批以知识融合为特征的情报项目,如Big Mechanism^[8](大机制项目)、Foresight and Understanding from Scientific Exposition^[9](根据科学论著进行预测与解读)项目、FOCUS^[10](开放环境中的反事实预测)项目和Mercury^[11](大数据安全预测)项目等。

本文分三个部分,首先总结当前大数据环境下情报学研究与实践知识融合的现状和趋势,梳理情报学研究与实践有关知识融合应用的特点;其次概述现有知识科学领域中知识融合研究与实践的内容及特征;最后根据以上两部分,概括总结情报学中知识融合的研究内容。笔者认为,大数据环境下,知识融合已经成为情报学的一个重要发展方向。一方面,已有的知识融合研究成果,对大数据环境下情报学

和情报工作的创新发展,起到了促进和支撑作用;另一方面,大数据环境对情报学和情报工作提出了新的要求,已有的知识融合研究成果还不足以帮助情报学和情报工作满足这些新的要求。为此,情报学正在结合自身的特点,探索大数据环境下情报领域中的知识融合理论和方法,用它们指导情报学研究和实践,并且取得了一些成果;同时进一步加强情报领域知识融合研究和实践的总结,将有可能形成情报学的知识融合范式和理论体系,从而促进情报学的发展。

1 情报学研究与实践知识融合的现状与趋势

1.1 情报理念:从辅助决策支持转向直接的预测和预警

传统的情报学强调情报是“耳目、尖兵、参谋”,起的是辅助决策的作用。大数据和人工智能时代的到来,情报环境发生了根本性的转变,兰德公司曾指出现代情报及情报工作面临复杂多变、深度不确定性的挑战^[12]。正因为如此,现代情报学越来越重视不确定情报环境下的预测和预警。换句话说,情报工作不再是二线工作,而是直接参与到问题解决和复杂决策之中,这与以往的情报工作“是科学研究的前期劳动”这样的定位和理念是完全不同的。多源信息融合、多模型分析的交叉验证、多假设的分析和比对参照、综合考虑多种影响研究对象的因素等逐渐成为现代情报工作的核心理念。现代情报工作已经不再是单纯的信息整序、简单的信息内容揭示和描述,情报研究成果也不再是对现有信息的简单罗列和简单统计或简单可视化展示,也不仅仅是对相关研究内容的各种观点和结论的单纯综述。现代情报工作更为看重的是,在全面掌握现有信息的基础上,运用科学的方法,对事物的发展趋势做出判断或对可能出现的问题提前给出提醒,并提出相应的解决方案。

在为决策提供辅助支持转向直接的预测和预警的过程中,知识融合功不可没,正是因为当

代情报工作中大量地应用了知识融合的理念和技术,并将它们与情报领域多年积累的思想和方法相结合,才使得情报工作在预测和预警方面取得了显著的成果。在理论上,相关学者已经证明了基于多源信息、知识融合和群体智慧的、由信息情报驱动的预测,是一种非常有效的预测方式。例如,美国著名预测专家菲利普·泰洛克(Philip E. Tetlock)经过十年的跟踪研究,发现领域专家由于过度自信而导致在预测中常常发生认知偏差,其预测的准确性并不比普通人高,而普通人组成的预测团队在经过适当的预测训练之后会采取更加谨慎的态度,综合多方面的信息和融合多种预测方法,通过融合多人的预测结果,准确率也非常高,由此,Tetlock提出了一种融合群体智慧的预测思想和方法体系,强调从多种来源收集证据,注重概率思维,组建团队一起预测,记录预测得分,愿意承认错误并及时调整预测结果^[13-14]。在实践上,知识融合已经被广泛地应用到情报工作中并卓有成效。美国国家研究委员会(NRC)下属的“未来颠覆性技术预测委员会”(CFDDT)通过大量的调研,发现颠覆性技术预测不仅要识别出潜在颠覆性技术的征兆信号,还要综合考虑市场、投资、成本、社会、人口、地理、政治、文化等多方面影响因素,此外,由于颠覆性技术本身所具有的复杂性,只用单一的预测方法、预测工具,预测团队并不能做出准确的预测。为此,该委员会进一步提出了一种用于预测颠覆性技术的理想方法论和框架,其特点是融合广泛来源的数据,融合多种预测方法,融合多样化团队的预测结果、多种数据处理工具,并在预测过程中能够不断持续地融合新的情报和信号^[15]。

1.2 情报采集:从传统的面向任务的被动式采集转变成面向内容理解的主动式感知

在情报采集方面,传统的情报采集通常是面向任务的被动式采集,即以任务的主题或内容要求为依据,采用人工、自动或人工与自动相结合的方式搜集与任务相关的情报信息。美国

国家情报总监办公室(ODNI)的大卫·艾萨克森(David Isaacson)曾指出,面向任务的被动情报采集方式在效率和可靠性方面存在瓶颈,难以应对新兴情报的挑战^[16]。大数据环境下,随着数据处理技术和智能算法的发展,情报采集的方法已经从传统的面向任务的采集向更加智能的情报感知转型^[17]。情报感知是获得对世界相关部分的必要理解的方法,它有意识地融合了与问题相关的知识^[18]。

一方面,面向内容理解的主动式感知在采集阶段就借助于情感分析、隐喻识别、实体识别、关系抽取等知识融合技术实现对信息内容的深度理解与信息场景的感知与识别,从而更加准确地过滤出与情报任务直接相关的有价值的信息,减少所采集到的信息中的“噪音”,避免信息“过载”。例如,Kejriwal等人设计了一种综合运用信息抽取、语义建模、自动聚类等技术融合技术的信息感知方法并进行了原型实现,从海量的Web信息源中发现并获取有价值的犯罪线索,为犯罪组织或成员建立了更加全面的画像及其犯罪活动踪迹,从而帮助执法机构打击人口贩运、武器走私等组织犯罪^[19];Szekely等人提出了一种基于大数据技术的面向领域Web的搜索、索引和分析系统,从多源、异构,海量的Web页中提取关键实体及关系信息并利用语义技术构建领域知识图谱,其特点之一是支持连续的、细粒度的网页信息感知^[20]。另一方面,也是更为重要的,面向内容理解的主动式感知可以根据情报任务的背景和已经采集信息所涉及到的场景,对信息采集的方向、范围等进行有意识的“优化”,甚至主动发现表面上并不相关或者在情报采集规划阶段并没有考虑到的信息。例如,美国高级研究和开发活动机构(Advanced Research and Development Activity)的NIMD(Novel Intelligence from Massive Data,从海量数据中获取新情报)项目^[21]中就包括了帮助情报分析人员发现、总结面向特定情报任务的最优信息搜索(采集)行为模式,并通过交互协助情报人员以最优方式完成精准高效的情报感

知。作为 NIMD 项目的一部分,斯坦福大学的 Fikes 等人设计和实现了一个名为 KANI (Knowledge Associates for Novel Intelligence) 的自动化关联系统^[22],将来自 Web 上多个数据源的知识进行关联聚合,通过知识推理(知识模式的匹配)进一步检查不同来源中有关同一对象的信息内容的逻辑一致性,如果发现存在逻辑上的不一致,系统则会重新考虑这些信息源的可靠性和真实性,从而调整信息采集的策略。

1.3 情报分析:从注重关联关系的挖掘进一步扩展到对因果关系的探究

近年来,情报学领域已经意识到,情报分析所使用的基于统计的数据挖掘方法,其主要功能是发现数据信息之间隐含的关联关系。虽然发现关联关系在一定程度上有助于问题的解决和促进事物的发展,例如,沃尔玛超市发现的啤酒和尿布销量之间的关联就是一个成功的范例。但是,仅有关联关系并不能有效地解释事物之间的因果及其背后所蕴含的机制。图灵奖得主朱迪亚·珀尔(Judea Pearl)指出,发现事物之间的关联关系是认识世界最基础的一层,要具有更高层次的分析能力就需要发现和构建更高一层的因果关系^[23]。越来越多的情报专家也已经认识到,识别并揭示事物之间的因果关系是进行高质量情报分析的重要基础和必备条件^[24]。传统情报分析中对因果关系的发现,主要是依靠情报分析人员的智慧来完成的,情报人员自身的经验和背景知识对分析的结果有很大影响。大数据环境下,情报领域中对因果分析的要求日趋强烈,正逐渐成为情报分析的“常态”。但是,大数据具有数据量大、结构复杂、形式多样的特点,其中往往会包含数量巨大的相关的和不相关的、显性的和隐性的多种因素,仅仅对这些因素加以甄别就有着巨大的工作量,更何况大数据中隐藏的关系结构一般都比较复杂、因果链冗长,一种因果关系需要多种类型数据的“集成”才能得以显露,或者需要分散在不同数据中的多种关系的“合成”才能得以发现。为

了充分弥补传统情报分析中因果分析及其方法的短板,满足大数据环境下情报分析注重因果关系发现这一转型要求,近年来情报界越来越重视情报分析中的知识融合,并且有意识地构建基于知识融合的情报分析理论和技术方法。

例如,美国国防部高级研究计划局(DARPA)于2014年提出了一项名为“大机制”(Big Mechanism)的项目,该项目认为,复杂系统中的因果关系是由分散在系统之中的众多“小片段”(Small Fragments)组成的,文献中所反映出来的因果规律也是如此,因果关系的片段分散在巨大、零碎,甚至有时相互矛盾的文献和数据库中,因此有必要开发一种自动阅读文献(文摘和全文)技术,从中抽取已知的因果片段,并将它们组装成完整因果模型,以便根据这些模型进行推理,从而获得相应的结论^[25]。“大机制”项目的这种理念实际上就是一种知识融合。该项目的首个应用领域是癌症生物学,从有关文献中抽取出论及 Ras 基因家族的癌症生物信息传导途径和机理,进而将它们融汇成大型的因果模型,利用这些模型识别出并解释大量先前并不明确的癌症信息传导途径^[26]。

1.4 情报服务:从知识服务走向智慧服务

大数据、人工智能技术与情报服务工作的融合,将形成全链条的智能化情报服务,这将是情报服务的发展方向^[27]。国外学者 Yeh 通过大量案例说明了在未来情报服务中人工智能将扮演重要的角色,指出了情报的智慧服务是未来情报服务发展的方向^[28];国内学者认为,智慧服务是继文献服务、信息服务、知识服务后又一新的发展阶段,并将成为情报服务的主流理念与方式^[29]。笔者认为情报服务将进入智慧服务时代,并且知识融合在智慧情报服务中起着关键作用^[30]。

智慧情报服务除了在学术界有广泛的研究以外,还得到了产业界和情报实践领域的重视。从国外的实践来看,近年来智慧情报服务中的深层次知识融合体现出以下三个方面的特点。

①应用智能信息技术是智慧服务的核心,与以

往信息技术的应用不同,智慧情报服务的主体正在从以人为主、机器为辅向机器为主、人机深度融合的方向发展,并且这种服务的外在表现与人类提供的服务基本相同。例如,美军指挥官在指挥过程中必须与众多参谋打交道并使用不同的计算机系统,需要在头脑中融合大量数据信息才能做出决策,这使得指挥官负担很重。有鉴于此,美国陆军通信电子研究与开发中心于2016年启动了“指挥官虚拟参谋”(Commander's Virtual Staff)项目^[31],该项目借鉴苹果的Siri、谷歌助理(Google Now)、IBM的Watson系统能够以自然的方式与用户交互,可以解决复杂的问题并能够持续地对环境和用户做更进一步的理解,旨在开发一套能够对来自各个方面的情报信息进行融合和智能处理、向使用者提供主动建议和高级情报分析结果,并可以针对使用者的独特需求和偏好与之进行自然交互的系统,以减轻指挥官的认知负担。②智慧情报服务在本质上服务于用户所关心的智慧问题,而不是知识问题,因此,智慧情报服务不仅要情报环境进行全面、全局的感知,还要通过融合背景知识和相关问题的隐性知识来深度理解情报问题和情报环境,只有如此,才能为服务对象提供知识性材料中所蕴含的深层次含义、意图等高阶认知性情报,从而将所提供的情报转化为用户的聪明才智,提升用户解决问题的能力。例如,DARPA战略技术办公室(STO)设立的COMPASS项目(Collection and Monitoring via Planning for Active Situational Scenarios)^[32],旨在研发一种能够帮助情报人员了解在战争“灰色地带”冲突中敌方真实意图的智慧情报服务系统。“灰色地带”冲突一般是指国家或地区之间发生的低于常规战争水平的冲突,这类冲突往往信息缺失、发展缓慢、对抗不一定激烈、手段多样、真正的诉求比较隐蔽,等等。COMPASS项目通过主动收集快速变化的冲突场景信息,深度融合背景知识(如当地文化、习俗、政治、经济、安全以及社会冲突等情况),不断解读“灰色地带”冲突背后的多种含义,最终确定对手的行动和意

图。③借助知识融合使所提供的情报产品具有可解释性是情报服务的重要趋势。目前情报智慧服务过程中所使用的人工智能技术,如深度学习,虽然可以拟合高度复杂的数据并且有较好的效果,但其内部推导过程是一个黑箱,这就使得利用这样的技术所产生的情报成果可解释性较差,用户只能知其然而不知其所以然。可解释是情报服务中非常重要的要素,单纯的没有解释说明的结果本身是不可靠的^[33]。因此,确保情报产品的可解释性,是智慧情报研究和实践中的核心课题。例如,施乐公司的帕洛阿尔托研究中心(PARC)开发了一种具有可解释性的智慧服务系统COGLE^[34],机器通过融合人类的知识(如人类学习的课程、操作手册、操作规范等)从而形成与人类相同的共识(Common Ground),利用这些共识,COGLE系统通过可视化交互界面为用户提供情报产品(结果)的解释和说明。

2 现有知识融合研究

从以上的介绍中可以看出,知识融合在情报活动中发挥着越来越重要并且是不可替代的作用,已经成为当代情报学和情报工作发展的重要推手和研究内容。知识融合研究发端于20世纪70年代,相关研究分布于不同的学科,其中与情报活动关系最为密切的知识融合主要是知识科学领域的研究和实践,其特点是研究对象不再限于用传感器获得的信息,而是研究已有信息库、知识库中知识之间的融合。下面先简要回顾知识科学视角下知识融合研究的发展历程,进而总结情报学中需要研究的知识融合的内容。

由于受制于不同时期认识水平和科技发展水平,知识融合的应用背景、研究的侧重点各不相同,本文根据主要研究内容以及代表性成果将知识融合的研究划分为三种类型,这三种类型也代表了知识融合的三个发展阶段,分别为基于代理的知识融合、基于模式的知识融合、基于机器学习的知识融合(见表1)。

表 1 知识融合的发展阶段

发展阶段	基于代理的知识融合	基于模式的知识融合	基于机器学习的知识融合
主要解决的问题	解决分布式环境下的知识共享、再利用和转化问题	解决面向特定领域上下文感知与决策问题	解决面向开放数据的大规模知识库自动组织与构建问题
知识融合方法	代理技术	本体语言、融合模式	机器学习
代表性成果	KSE、KRAFT	Smimov 融合模式	NELL、Knowledge Vault
主要应用领域	专家系统	军事指挥、决策支持系统	互联网知识服务

2.1 基于代理的知识融合

基于代理的知识融合是知识融合研究的早期代表,其核心任务是解决分布式信息环境下知识的共享、再利用和转化问题,通过中间件代理实现对众多知识资源的搜索和抽取,并对异质知识进行转化与集成形成统一的视图,为专家系统更高效地解答问题提供支持。

早期知识融合主要应用于专家系统,专家系统一般主要由知识库与推理引擎两部分组成,当时的知识库都是由相关领域专家人工构建的,具有很高的准确率和利用价值,但是其构建过程耗时耗力而且存在覆盖性较低的问题。为了解决这个问题,就要实现多领域或学科知识库之间的知识共享、再利用和融合。为了将分布式环境下的独立知识库集成到一起,同时又不破坏原有知识库的结构,最早的知识融合系统使用了两种代理:一种是封装器代理(Wrapper),它用于封装来源知识库并对外提供服务;一种是中介器代理(Mediator),它用于选择和执行对来源知识库的查询并融合特定的知识约束,从而为用户提供统一的知识视图。考虑到代理技术在知识融合方面所具有的优势,美国国防部高级研究计划局(DARPA)和美国国家科学基金会(NSF)在20世纪末联合设立了“知识共享计划”(Knowledge Sharing Effort, KSE)^[35],旨在开发有助于共享和重用知识库和基于专家系统的规范和组件,其中开发的 KQML (Knowledge Query and Manipulation Language, 知识查询和处理语言)是一种被广泛应用的交换信息和知识的语言和协议,用于支持分布式环

境下代理之间的实时知识共享和协同处理。由英国工程和自然科学研究委员会(EPSRC)资助的 KRAFT (Knowledge Reuse and Fusion/Transformation)^[36]项目进一步提出了基于多代理技术的知识融合机制和体系架构,KRAFT 所有部件都以代理的形式实现,支持对分布式信息环境下异构数据库中的数据和知识库中的知识进行融合。多代理知识融合系统的另一个典型代表是欧盟空间研究与开发办公室与俄罗斯圣彼得堡信息学与自动化研究所共同开展的可扩展信息空间的快速知识融合(Rapid Knowledge Fusion in the Scalable Infosphere)项目^[37],该项目进一步优化了 KRAFT 项目的知识融合代理架构,使知识融合系统更加具有可灵活性、可定制性和可伸缩性。

2.2 基于模式的知识融合

基于模式的知识融合的主要目标是解决复杂场景下的上下文态势感知与决策问题。21 世纪初语义网技术(Semantic Web)特别是本体语言及技术的出现,极大地提升了对复杂知识的表征能力,本体语言可以将知识的内部结构和外部结构进行形式化的描述和表示(如概念、属性、关系、域、实例等),知识融合主要是通过对多源的本体表示的知识的内外部结构、属性的改变、重组而实现知识之间的融合,在这个过程中,同类或类似问题域中的形式化知识的融合,有一定的相似性,寻找或发现这种相似性,并将它们总结成融合模式,就成为了这一阶段知识融合研究的核心内容。所谓融合模式是知识融

合系统或知识服务在解决专有领域问题时,在不同阶段所遵循的流程和规则的形式化抽象,它实际上就是从实践中总结和抽象出来的知识融合的规则和规律,这些规则和规律可以在解决相似问题时重用,从而提高问题求解的效率。

基于模式的知识融合在情境感知以及在涉及有效决策的相关任务中体现出了巨大的应用潜力和优势。例如,态势感知需要了解上下文和解释感兴趣的事件和行为,这就需要利用融合模式来推断相似环境中与特定目标、特定能力相关的关键特征以及决策者通常考虑的政策等^[38]。例如,Scherl等^[39]针对美军作战指挥需要提出了四种知识融合模式,分别为战场运动分析融合模式、国土安全问题融合模式、联合作战互操作融合模式、战场替代方案融合模式;Little等人^[38]提出了灾后救援等涉及复杂的事件、对象、行为等动态态势感知的高级知识融合模式,该模式通过捕获构成复杂状态和威胁的各种复杂关系、实体和事件来帮助理解与时间变化、上下文信息和身份相关的态势,用以生成特定领域的情景假设以及对这些情景假设进行高级推理和融合;Smirnov等人^[40]对已有的知识融合模式进行广泛的调研,归纳总结出了7种知识融合模式:简单融合、扩展融合、实例化融合、配置融合、自适应融合、平面融合和历史融合,并对7种融合模式的应用场景和产生结果进行了讨论。此后,Smirnov等人进一步针对“到目前为止,尚未开发出通用的知识融合框架、通用的知识融合技术、算法或过程”,对知识融合的模式进行了概括和总结,提出了三大类27种融合模式,以便用这些找到的知识融合模式“设计出基于知识场景感知系统的通用框架”^[41]。

2.3 基于机器学习的知识融合

前两个阶段的知识融合主要集中于为专家系统或知识融合系统中的问题求解提供支撑,基于机器学习的知识融合则主要通过机器学习等自动化方法实现对开放数据的自动知识抽取与学习、知识链接的建立、知识的统一等,最终

实现对大规模知识库的自动构建和知识自动组织。以HDFS、Hadoop、Pregel为代表的大数据技术、图计算技术的兴起,解决了大规模知识计算与存储的问题,这为构建统一大规模的知识库提供了技术支撑。Web 2.0时代的到来,指数级增长的信息使网络中蕴含了大量的常识性知识,相继出现了许多开放域链接数据库如DBpedia、Wikipedia、Freebase等,为知识库的大规模构建提供了新的途径。同时面向Web的开放文本抽取技术也得到了进一步的发展,为大规模自动知识库的构建提供了方法支撑。在这种情况下,学者开始转向解决大规模知识库自动构建过程中的融合问题,即如何将Web中学习到的新知识更新融入到原有知识库中。

基于机器学习的知识融合的典型实例有:卡内基梅隆大学的NELL系统^[42-43]、佛罗里达大学的SigmaKB系统^[44]以及Google公司开发的Knowledge Vault知识库^[45]。NELL系统可以通过机器学习算法24小时不间断地从海量的Web信息中抽取事实型知识并不断地扩充和更新知识库,而且能保证知识库中知识的正确性和一致性。SigmaKB系统则可以利用共识最大化融合方法实现对多达79个知识库的检索,并为最终用户提供高质量的知识服务。Knowledge Vault是一个面向Web数据的超大规模的概率知识库,其最大的特点是采用监督式机器学习方法来融合海量的信息源,通过概念推理系统完成事实的构建与关联,并最终将新知识更新到统一的知识库/知识图谱中。

2.4 小结

综上所述,我们可以看到,时代背景不同,对知识融合要求也不一样,知识融合的内涵、研究的内容以及主要解决的问题也各不相同,从总体上看,现有知识融合的研究有以下特点。

(1) 现有知识融合的核心是关注知识融合的具体实现,主要是如何设计与实现一个高效的计算机化知识融合系统,也就是说,用智能软件来实现人的知识集成、知识推理、知识服务等

一直是研究人员追求的目标。在这方面,知识融合系统的研究与情报活动有部分的重叠,但两者又有不同,知识融合系统研究追求的是算法和结构的最优,而情报活动中的知识融合研究并不限于系统实现和算法优化,其范围更为广泛,更强调对情报工作的适用性和实用性。

(2) 现有的知识融合已经开始引入认知科学、心理科学、知识管理等学科的思维和视角,探讨知识融合的机理机制、融合模式等原理性问题,试图为知识融合提供理论基础与程序化的路径,例如 Smirnov 等人的研究就是这方面的典型代表。不过,如前所述,由于现有知识融合的核心是系统实现,这些研究归根到底还是为知识融合系统的开发服务的,而情报活动是一种社会活动,从情报的采集到情报服务,人的因素始终贯穿其中,其中的知识融合本身就是认知问题、心理问题或知识管理问题,也就是说,情报学中的知识融合研究课题,大部分蕴含了知识融合的一般规律,其研究恰恰能为知识科学视角下的知识融合提供最为需要的理论基础。

(3) 现有知识融合中有关技术方法的研究,例如知识表示、知识融合算法、知识挖掘方法、知识图谱技术、知识库技术等,为情报活动中的知识融合提供了底层技术支撑,但不能直接解决情报活动中所有知识融合问题。情报学中的知识融合研究还需要将这些底层技术方法与具体情报实践相结合,根据情报需求和情报任务的不同,或者在某种底层技术方法的基础上进行参数调整,或者将多种底层技术方法加以“组合”和“集成”,使它们相互配合解决情报问题,甚至在情报活动的知识融合实践中创造出全新的技术方法。

从上述特点看出,现有的知识融合不能替代情报领域中的知识融合研究,本文第一部分谈到的知识融合问题,有些属于这个范畴,有些已经远远超出了现有研究的范围。例如,情报领域也涉及了态势感知和决策支持,但是情报领域的态势感知更多面向的是开放、复杂、不确定的环境,

而且知识融合要求的抽象层次更高,而现有的知识融合要求问题的定义要明确、边界要清晰,输入输出的知识是比较确定的。再比如,情报分析属于高级别的认知活动,会涉及复杂的、内隐的知识融合活动,而现有的知识融合研究还无法实现高级别情报分析的知识融合。由此可见,我们需要根据情报学 and 情报工作的发展和实践,系统总结知识融合的研究框架。

3 知识融合——情报学的一个重要发展方向

如本文第一部分所述,情报领域对知识融合的需求和应用越来越明显,知识融合已经成为推动未来情报学与情报工作发展的重要要素。其实,情报学本来就是与知识研究有非常密切关系的学科,学界也普遍认为,知识是情报学研究的一种重要对象^[46]。近年来,情报学和情报工作之所以有如此的长足发展,应该说,一方面是因为情报学和情报工作广泛地汲取了知识科学领域中有关知识融合的研究成果,特别是借鉴和应用了知识科学领域知识融合的技术和方法,并在情报实践中进行了适应性改造;另一方面,知识科学领域的知识融合研究并不能为所有的情报问题提供解决方案,情报领域的许多知识融合问题都需要情报研究者和情报实践者自己来解答,本文第一部分中的实例已经充分说明了这一点。可以确信,随着情报研究者的努力和情报实践的进展,知识融合必将在情报工作中发挥更大的作用,并成为情报学研究的一个重要热点领域,甚至成为情报学研究与实践的一种新范式。

为了更好地揭示情报学研究这一发展趋势,结合近年来的情报实践,我们概括出一个知识融合研究的总体框架(见图1)。一方面,这个框架对现有情报学研究和实践进行概括和总结,反映当前情报学的新进展、新特征,另一方面,试图勾画出未来情报学的一种发展路径,从而推动情报学的发展。

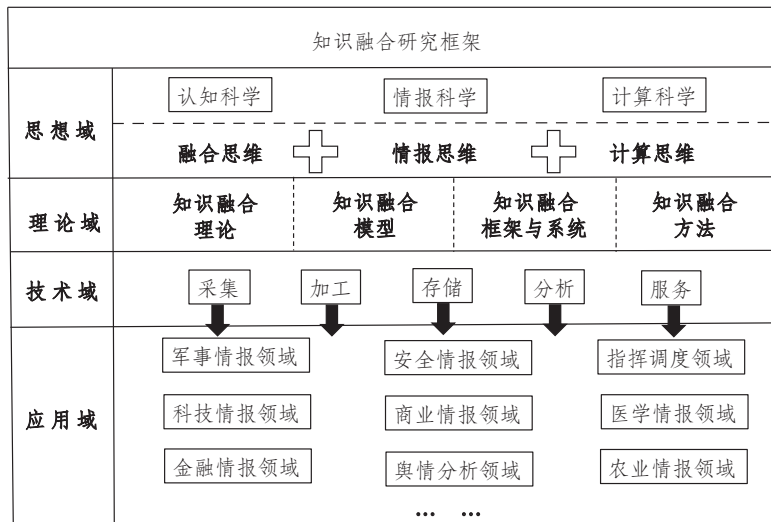


图 1 知识融合的研究框架

3.1 知识融合的思想域

思想域是应对复杂、多变、深度不确定情报环境的一组指导思想,在思想上、理论上指导、引领情报研究和情报工作。情报领域知识融合的思想域由融合思维、情报思维、计算思维共同组成,三种思维相辅相成。这三种思维的主要基础学科或来源学科分别是认知科学、情报科学和计算科学。

融合思维是开展现代情报工作最为核心的方法论和基础,情报工作特别是大数据环境下的情报工作需要融合思维的指导,从而有效应对不确定性极强的情报环境。近年来的实践表明,一方面,认知科学对指导情报工作中的知识融合发挥着重要的指导作用,认知科学从心智的角度探究人脑知识的处理与加工、知识的表征、知识建构、知识推理、知识应用等内容,为认识情报工作中知识融合的机制和原理提供了有力的支撑和依据^[47],另一方面,认知科学的方法和工具已经在情报工作知识融合的实现过程中发挥了重要作用^[48]。目前,关于融合思维研究还比较少,多是隐含在具体情报问题的研究之中以及情报实践的项目里,还需要我们对此做深入的挖掘,特别是需要从更广泛的视野即认知角度来总结提炼情报工作中的融合思维,以

便指导情报工作者有意识地用融合的理念来思考 and 解决情报问题。

情报思维是在情报活动中对所蕴含的一般性的、简约化的行动原则和方法体系的总结与提炼,它体现了情报工作的特殊性,是从事情报活动的核心指导思想。我国学者在 20 世纪 80 年代中期就开始对情报思维进行讨论,但没有形成规模,研究的内容主要包括情报思维的概念^[49]、情报思维的基本要素结构和组成^[50]、情报思维的结构^[51]、情报思维的类型及其在情报活动中的作用^[52]等,这些研究极大地推进了情报思维的研究。但是,到目前为止,情报界还没有对情报思维的具体内容达成共识,特别是没有总结出精炼的、情报特色鲜明的情报思维的具体表述。关于情报思维仍然是值得我们继续探索的课题。

与上述融合思维、情报思维相比,计算思维的研究最为成熟,其核心内容也被学界广为接受。所谓计算思维就是运用计算的概念去解决问题、设计系统以及理解人类行为的分析方式^[53],或者是通过人类对物理世界及其层次、问题的抽象(Abstraction),采取自动化(Automation)的方式让“计算机”(可以是实体机器或是支持人机交互的计算机)去处理、解析这些抽象^[54]。简

单地说,计算思维的核心内容就是模型化、形式化、自动化。作为一种解决问题的思考方式,计算思维在解决情报问题时,首先抽象出实际问题的组成要素及其关系,并用简单明了的方式表达出来,即模型化;其次,将模型中的要素和关系用特定的数学符号代替和表达,即形式化;最后,把解决问题的思路变成计算机指令,由机器执行并得出结果,即自动化。研究和实践表明,计算思维在指导解决融合问题和情报问题方面发挥了重要的作用,起到了非常好的效果。当前情报领域应该重点研究如何在情报活动中有意识地应用情报思维,并对情报思维在情报活动中的实现规律进行深入的研究。

3.2 知识融合的理论域

知识融合的理论域主要包括四个方面的内容,分别为知识融合理论、知识融合框架与系统、知识融合模型、知识融合方法。

(1) 知识融合理论研究

理论是指人类对自然、社会现象按照已有的实证知识、经验、事实、法则、认知以及经过验证的假说,经由一般化与演绎推理等方法,进行合乎逻辑的推论性总结^[55]。目前有关知识融合的理论研究散见于诸如知识工程、教育学、心理学、工程技术等不同的相关学科,现有的理论通常与学科紧密相连,彼此之间不够统一,有一定的差异,学科之间的相互借鉴还有一定的障碍。情报领域的知识融合除了要借鉴这些研究以外,还要针对情报活动的特点,对知识融合的一些基本问题做进一步的深入研究,这些基本问题包括:①知识融合基本概念,即情报领域知识融合的内涵和外延是什么,知识融合具备哪些基本的特性,情报领域中知识融合与“知识组合”“知识整合”“知识汇聚”“知识迁移”等概念的区别与联系是什么;②情报领域知识融合的研究目的,即大数据时代知识融合研究要解决的核心问题是什么;③构成情报认知的知识融合的原理是什么,不同类型知识融合的机制是什么,知识融合会产生什么样的效应,等等。

(2) 知识融合模型研究

模型是计算思维中“抽象”和“表示”的具体体现。抽象的本质是将要解决的问题进行表示和建模,建模的过程就是理解问题、解决问题的过程^[29],从计算和融合的角度看,情报领域的知识融合模型首先是一种数学模型,用数学方法描述和刻画;其次它能够通过机器进行表示、模拟与计算,进而获取通过融合所产生的结果。知识融合模型研究包括情报与知识表示的建模与融合方式的建模。情报的知识融合模型是未来情报领域的重要研究内容。从近年的情报实践看,自动建模将是未来的一个发展趋势,例如,Lippmann 等人^[56]开发了一种数据驱动模型发现框架,该框架能够借助强化学习技术根据输入的数据集自动搜索最优模型并进行参数的调整,从而帮助那些有领域专业知识但没有数学和计算机知识背景的情报分析人员快速构建真实、复杂过程的情报分析模型。这项研究对情报领域的知识融合模型研究有非常大的启迪意义,指出了情报领域知识融合模型研究的未来发展方向。

(3) 知识融合框架与系统研究

知识融合框架本身并不解决具体的知识融合应用问题,它只是为实现融合计算而给出的一个用于插接、组合解决问题的相关组件的基础。知识融合框架是在解决决策、预测、干预等问题的过程中为知识融合而设计的具有一定约束性的支撑结构,由两层组成:①用于知识融合业务和逻辑表示的抽象层,表现为一组抽象构件及构件实例(信息、知识、模型、算法)间交互的方法框架,还表现为框架内对象间的交互模式和控制流模式^[57]。②用于实现知识融合过程的计算层,一方面,它界定了实现知识融合的技术边界,进而将相关的软件组件约束在这个边界内,另一方面,它还用来提供支撑知识融合的可选的配套软件组件或工具,可以根据具体知识融合应用来扩展、安插更多的组成部分,从而更迅速和方便地构建完整的面向领域的知识融合系统或方案。例如,前文提到的美国未来颠覆性技术预测委员会(CFFDT)所提出的用于颠覆性

技术预测的融合框架就是这方面研究的典型代表,该框架在数据层面提供多源数据收集和融合工具,在预测过程中能够不断融合新的情报和信号,在预测方法上融合了多种预测方式。

(4) 知识融合方法研究

方法是指为达成某个特定目的,可以用来实践(实际操作)的模式或过程(步骤),也包括使用的工具或技巧。知识融合的方法具有可计算特性,包括知识融合的模式和知识融合的通用算法。知识融合模式是在知识融合某类应用中所使用到的一般性和普适性的规律和操作过程。例如上文中提到的 Smirnov 等人针对不同的知识融合应用场景总结出七种典型的知识融合模式^[40],这类模式可以移植到相同或相似的应用场景中使用。知识融合的算法是在情报思维和融合思维的指导下,以通用的算法设计、实现的思路为基础所构建的一系列计算步骤和指令序列。例如, Karpatne 等人提出了一种融合知识驱动和数据驱动的算法,能够将先验知识应用于数据密集型的情报分析任务中以解决分析数据过于稀疏的问题^[58]。

3.3 知识融合的技术域

情报工作包括多个环节,如采集、加工、存储、分析、服务等,当然,这种划分是比较粗粒度的,每一个环节还可以再进一步细分,或者可用不同的标准划分出不同的环节,图1技术域中显示出的环节仅仅是示意性的,旨在说明需要根据情报工作环节的特征来开展知识融合技术研究。情报工作每一个环节要解决的问题不同,所需要的核心技术或主要技术也会有所不同。以情报采集环节中情报感知为例,情报感知是情报活动中最基础也是最重要的认知单元,一般研究如何通过多源信息、知识的融合形成对情报环境全局态势的理解与意义的构建,在这个过程中,需要对信息进行交叉验证、多源信息融汇、信息的语义理解以及信息的再组织与存储,等等,这就需要对原有的情报采集技术进行改进,增强其融合功能,使之具备情报感知

的能力。换句话说,情报领域的知识融合技术研究,需要根据情报工作各环节的需求、特征来开展。具体而言,知识融合技术的内容包括但不限于以下内容:①情报活动各环节中知识融合的特征规律与实现技术;②情报活动各个环节中的知识融合技术与已有技术的互操作及融会贯通;③情报活动各环节的技术之间的逻辑关系及相互作用机理,形成情报活动的知识融合技术体系;④为适应不同情报需求和任务的多种融合技术的集成应用及其实现;⑤对现有知识融合技术方法的适应性改造,提升现有技术方法对情报活动的适用性和可操作性;⑥各种知识融合技术需要的基础资源建设(例如轻量级本体的建设与实现)等。

3.4 知识融合的应用域

近年来,随着社会、经济、技术的不断发展,情报受到了空前的重视,一方面情报学领域的情报思维、情报方法、情报工具被其他领域和行业所借鉴和使用,另一方面情报学及其学者也在不断扩展新的研究对象,从而产生了与该领域高度相关的应用分支领域,包括军事情报领域、科技情报领域、金融情报领域、安全情报领域、商业情报领域、舆情分析领域、指挥调度领域、医学情报领域、农业情报领域等。这些领域的情报活动往往会涉及专业知识,属于专业密集型的应用。由于不同应用领域的行业特性和情报环境千差万别,这些领域中的知识融合应用也会存在很大的差异,这就要求研究人员在情报活动的一般性知识融合理论和方法的基础上,根据应用领域和应用场景特点,研究具有领域特色的知识融合技术方法体系。例如,金融情报领域更加注重金融事件的预测与预警,这对于金融机构把控风险、及时发现投资机会都是至关重要的。金融情报的数据既包括数值型的数据,又包括来自市场外部的文本型数据,既包括基于时序的线性数据,也包括大量的非线性数据,这就要求结合金融情报活动的特点,结合金融学的知识,研究金融情报活动中的知识

融合模型、算法,甚至需要对知识融合的计算框架进行细化和再设计。又例如,在指挥调度领域,全面、深入的情报感知,高效、准确、及时的决策是该领域关注的重点。特别是在应急指挥领域,准确、全面的情报感知对于做出正确的决策具有非常重要的意义,这就要求该领域的知识融合研究要能够从情报感知中提取并抽象出准确的态势模式,以精确反映事件的走向,帮助决策者做出正确的决策。

4 结语

大数据和人工智能时代,情报学和情报工

作正发生着翻天覆地的变化,知识融合正在其中发挥着巨大的潜能和优势。知识融合不仅仅是技术,更是理论和方法,是我们在面对复杂、多变、不确定世界时认知世界、解决问题、做出正确决策的有效思想和路径。情报学作为一门研究知识的学科,有必要充分利用自身已有的积淀和优势,深入研究、总结、发现知识融合在情报活动中的机制机理以及方法和技术。近年来国内外的实践表明,情报学的研究者和实践者已经在这方面进行了有益尝试,取得了一些有价值的成果,反映了情报学的一种新的发展趋势,笔者相信,这一新趋势必将进一步促进情报学科体系的完善和发展。

参考文献

- [1] MacDonald M S, Oettinger A G. Information overload; managing intelligence technologies[J]. Harvard International Review, 2002, 24(3): 44-48.
- [2] ZDNET. NSA is so overwhelmed with data, it's no longer effective, says whistleblower[EB/OL]. [2020-03-14]. <https://www.zdnet.com/article/nsa-whistleblower-overwhelmed-with-data-ineffective/>.
- [3] Brantly A F. When everything becomes intelligence; machine learning and the connected world[J]. Intelligence and National Security, 2018, 33(4): 562-573.
- [4] Schacher G, Irvine N, Hoyt R B. Joint intelligence operations centers(JIOC) business process model & capabilities evaluation methodology[R]. Monterey, California: Naval Postgraduate School, 2012.
- [5] 张心源, 邱均平. 大数据环境下的知识融合框架研究[J]. 图书馆学研究, 2016(8): 66-70, 43. (Zhang Xinyuan, Qiu Junping. Research on knowledge fusion framework in big data environment [J]. Research on Library Science, 2016(8): 66-70, 43.)
- [6] 唐晓波, 朱娟, 杨丰华. 大数据环境下的知识融合框架模型研究[J]. 图书馆学研究, 2016(1): 32-35, 18. (Tang Xiaobo, Zhu Juan, Yang Fenghua. Research on knowledge fusion framework model in big data environment [J]. Research on Library Science, 2016(1): 32-35, 18.)
- [7] 王曰芬, 岑咏华. 大数据时代知识融合体系架构设计研究[J]. 数字图书馆论坛, 2016(10): 16-24. (Wang Yuefen, Cen Yonghua. Research on the architecture design of knowledge fusion system in the era of big data [J]. Digital Library Forum, 2016(10): 16-24.)
- [8] Cohen P R. DARPA's big mechanism program[J]. Physical Biology, 2015, 12(4): 045008.
- [9] IARPA. Foresight and understanding from scientific exposition(FUSE)[EB/OL]. [2020-03-14]. https://www.iarpa.gov/index.php/working-with-iarpa/index.php?option=com_content&view=article&id=31&Itemid=177.
- [10] IARPA. Forecasting counterfactuals in uncontrolled settings(FOCUS)[EB/OL]. [2020-03-14]. <https://www.iarpa.gov/index.php/research-programs/focus>.
- [11] IARPA. Mercury[EB/OL]. [2020-03-14]. <https://www.iarpa.gov/index.php/research-programs/mercury>.
- [12] Lempert R J, Popper S W, Bankes S C. Shaping the next one hundred years; new methods for quantitative, long-term policy analysis[M]. Santa Monica, CA: Rand Corporation, 2003.
- [13] Tetlock P E, Gardner D. Superforecasting: the art and science of prediction[M]. New York, USA: Crown, 2016.
- [14] Tetlock P E. Expert political judgment; how good is it? How can we know?[M]. Princeton, USA: Princeton University Press, 2017.
- [15] National Research Council, Committee on Forecasting Future Disruptive Technologies. Persistent forecasting of dis-

- ruptive technologies—report 2[M]. Washington,DC:National Academies Press,2010.
- [16] National Academies of Sciences,Engineering,and Medicine. Challenges in machine generation of analytic products from multi-source data[M]. Washington,DC:National Academies Press,2017.
- [17] 李广建,罗立群. 计算型情报分析的进展[J]. 中国图书馆学报,2019,45(4):29-43. (Li Guangjian,Luo liqun. Progress in computational intelligence analysis [J]. Journal of Library Science in China,2019,45(4):29-43.)
- [18] Moore D T. Sensemaking:a structure for an intelligence revolution[M]. Government Printing Office,2013.
- [19] Kejriwal M,Szekely P,Knoblock C. Investigative knowledge discovery for combating illicit activities[J]. IEEE Intelligent Systems,2018(1):53-63.
- [20] Szekely P,Knoblock C A,Slepicka J,et al. Building and using a knowledge graph to combat human trafficking [C]//Proceedings of 14th International Semantic Web Conference. Bethlehem, PA, USA: Springer, 2015: 205-221.
- [21] SourceWatch. Novel intelligence from massive data [EB/OL]. [2020-03-14]. https://www.sourcewatch.org/index.php/Novel_Intelligence_from_Massive_Data.
- [22] Fikes R,Ferrucci D,Thurman D. Knowledge associates for novel intelligence (KANI) [C]//2005 International Conference on Intelligence Analysis. 2005.
- [23] Pearl J,Mackenzie D. The book of why:the new science of cause and effect[M]. New York,USA:Basic Books, 2018.
- [24] Stigler A L. Assessing Causality in a Complex Security Environment[J/OL]. Joint Force Quarterly,2015,76 [2020-03-14]. <https://ndupress.ndu.edu/JFQ/Joint-Force-Quarterly-76/Article/577586/assessing-causality-in-a-complex-security-environment/>.
- [25] DARPA. DARPA big mechanism[EB/OL]. [2020-03-14]. <https://www.darpa.mil/program/big-mechanism>.
- [26] Valenzuela-Escárcega M A,Babur O,Hahn-Powell G V,et al. Large-scale automated machine reading discovers new cancer-driving mechanisms[J]. Database (Oxford),2018:bay098.
- [27] 冯秋燕,朱学芳. 人工智能在情报工作中的应用研究[J]. 情报理论与实践,2019,42(11):27-33. (Feng Qiuyan,Zhu Xuefang. Application research of artificial intelligence in intelligence work [J]. Information Studies; Theory & Application,2019,42(11):27-33.)
- [28] Yeh P F. The case for using robots in intelligence analysis[J]. Studies in Intelligence,2015,59(4):1-8.
- [29] 黄幼菲. 图书馆知识服务的扬弃和“飞跃”:公共智慧服务[J]. 情报理论与实践,2013,36(2):26-30. (Huang Youfei. Subetion and “leap” of library knowledge service;public wisdom service [J]. Information Studies;Theory & Application,2013,36(2):26-30.)
- [30] 罗立群,李广建. 智慧情报服务与知识融合[J]. 情报资料工作,2019,40(2):87-94. (Luo Liqun,Li Guangjian. Smart intelligence service and knowledge fusion [J]. Information and Documentation Services,2019,40(2):87-94.)
- [31] US Army. Army applies computer automation to operational decision making[EB/OL]. [2020-03-14]. https://www.army.mil/article/148549/army_applies_computer_automation_to_operational_decision_making.
- [32] Barlos F,Skinner A,Peeke R,et al. Collection and monitoring via planning for active situational scenarios(COMPASS)[R]. Sandia National Lab. (SNL-NM),Albuquerque,NM(United States),2020.
- [33] Gunning D. Explainable artificial intelligence(xai) [EB/OL]. [2020-02-13]. <https://www.darpa.mil/attachments/XAIProgramUpdate.pdf>.
- [34] PARC. Explainable AI;an overview of PARC's COGLE project with DARPA[EB/OL]. [2020-03-14]. <https://www.parc.com/blog/explainable-ai-an-overview-of-parcs-cogle-project-with-darpa>.
- [35] Patil R S,Fikes R E,Patel-Schneider P F,et al. The DARPA knowledge sharing effort:progress report[C]//Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning. 1992.
- [36] Gray P M D,Preece A,Fiddian N J,et al. KRAFT:knowledge fusion from distributed databases and knowledge bases[C]//Proceedings of the 8th International Conference on Database and Expert Systems Applications. IEEE, 1997:682-691.
- [37] Smirnov A V. Rapid knowledge fusion into the scalable infosphere;a concept and possible manufacturing applications[C]//Proceedings of the International NAISO Congress on Information Science Innovations(ISI2001),Symposium on Intelligent Automated Manufacturing(IAM2001). Dubai,UAE,2001.

- [38] Little E G, Rogova G L. Designing ontologies for higher level fusion[J]. Information Fusion, 2009, 10(1): 70-82.
- [39] Scherl R, Ulery D L. Technologies for army knowledge fusion[R]. Department of Computer Science Monmouth University, 2011.
- [40] Smirnov A V, Levashova T, Shilov N. Patterns for context-based knowledge fusion in decision support systems[J]. Information Fusion, 2015, 21: 114-129.
- [41] Smirnov A, Levashova T. Knowledge fusion patterns: a survey[J]. Information Fusion, 2019, 52: 31-40.
- [42] Carlson A C, Betteridge J, Kisiel B, et al. Toward an architecture for never-ending language learning[C]//Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence. Atlanta, Georgia, USA, 2010: 11-15.
- [43] Mitchell T M, Cohen W, Hruschka E, et al. Never-ending learning[J]. Communications of the ACM, 2018, 61(5): 103-115.
- [44] Rodríguez M, Goldberg S, Wang D Z. SigmaKB: multiple probabilistic knowledge base fusion[J]. Proceedings of the VLDB Endowment, 2016, 9(13): 1577-1580.
- [45] Dong X, Gabrilovich E, Heitz G, et al. Knowledge Vault: a web-scale approach to probabilistic knowledge fusion[C]//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2014: 601-610.
- [46] 马费成. 情报学的进展与深化[J]. 情报学报, 1996(5): 337-343. (Ma Feicheng. Progress and deepening of information science [J]. Journal of the China Society for Scientific and Technical Information, 1996(5): 337-343.)
- [47] Wang Y. Formal cognitive models of data, information, knowledge, and intelligence[J]. WSEAS Transactions on Computers, 2015, 14(3): 770-781.
- [48] Glodek M, Honold F, Geier T, et al. Fusion paradigms in cognitive technical systems for human-computer interaction[J]. Neurocomputing, 2015, 161: 17-37.
- [49] 陈建龙. 论情报思维及其概念来源[J]. 情报学刊, 1993(5): 328-333. (Chen Jianlong. On intelligence thinking and its concept source [J]. Journal of the China Society for Scientific and Technical Information, 1993(5): 328-333.)
- [50] 洪峰. 再论情报思维[J]. 情报杂志, 1991(2): 60-63. (Hong Feng. On intelligence thinking [J]. Journal of Intelligence, 1991(2): 60-63.)
- [51] 应思德, 沈浩. 论情报思维: 完善与发展[J]. 情报杂志, 1990(3): 33-39. (Ying Side, Shen Hao. On intelligence thinking: improvement and development [J]. Journal of Intelligence, 1990(3): 33-39.)
- [52] 吴国恩. 论现代情报思维[J]. 图书情报工作, 1992(5): 5-9. (Wu Guoen. On modern intelligence thinking [J]. Library and Information Service, 1992(5): 5-9.)
- [53] Wing J M. Computational thinking[J]. Communications of the ACM, 2006, 49(3): 33-35.
- [54] Wing J M. Computational thinking and thinking about computing[J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2008, 366(1881): 3717-3725.
- [55] 维基百科. 理论[EB/OL]. [2020-03-14]. <https://zh.wikipedia.org/wiki/%E7%90%86%E8%AB%96>. (Wikipedia. Theory [EB/OL]. [2020-03-14]. <https://zh.wikipedia.org/wiki/%E7%90%86%E8%AB%96>.)
- [56] Lippmann R P, Campbell W M, Campbell J P. An overview of the DARPA data driven discovery of models(D3M) program[C]//Proceedings of the 29th Conference on Neural Information Processing Systems(NIPS 2016). Barcelona, Spain, 2016.
- [57] 陈晓红. 信息系统教程[M]. 北京: 清华大学出版社, 2003. (Chen Xiaohong. Course of information system [M]. Beijing: Tsinghua University Press, 2003.)
- [58] Karpatne A, Atluri G, Faghmous J H, et al. Theory-guided data science: a new paradigm for scientific discovery from data[J]. IEEE Transactions on Knowledge and Data Engineering, 2017, 29(10): 2318-2331.

李广建 北京大学信息管理系教授, 博士生导师。北京 100871。

罗立群 北京大学信息管理系博士后, 讲师。北京 100871。

(收稿日期: 2020-04-13)