

# 中国科学院文献情报中心“十四五”发展思考

## ——基于数据、信息、知识与情报的规划框架设计

刘细文

(中国科学院文献情报中心, 北京 100190)

**摘要:** 本文系统分析了新时期文献情报机构面临的挑战和发展趋势, 分析新时期文献信息机构的工作对象、内涵、理论和技术的变化, 阐明数字化科研模式变革与需求变化对文献情报工作的影响和借鉴意义, 将数据替代信息作为图书情报工作流程设计的起点。全面归纳中国科学院文献情报中心的近期工作进展, 包括文献信息保障能力、数据汇聚和平台管理能力、科研管理与科技创新服务能力、学术信息评价与交流等方面。“十四五”期间, 中国科学院文献情报中心将面向“战略决策一线、科技创新一线、区域与产业一线”, 以战略情报服务和学科情报服务为核心, 建设国际一流的国家科技知识服务中心。

**关键词:** 中国科学院文献情报中心; “十四五”; 发展规划

中图分类号: G251 DOI: 10.3772/j.issn.1673-2286.2021.05.003

**引文格式:** 刘细文. 中国科学院文献情报中心“十四五”发展思考——基于数据、信息、知识与情报的规划框架设计[J]. 数字图书馆论坛, 2021 (5): 12-16.

发展规划的制订, 需要考虑已有发展基础, 把握未来发展趋势, 明确下一步发展思路, 作出明确的发展道路选择。笔者从以下方面谈谈对中国科学院文献情报中心制订“十四五”发展战略的基本思考。

### 1 对变化趋势的几点认识

文献信息机构面临的变革是多方面的, 其影响也是颠覆性的。对于中国科学院文献情报中心这样的科研信息服务机构来说, 需要从服务需求、工作内涵、图情理论3个角度认知和响应。

第一, 科学研究模式已经发生实质性变化, 这样一些变化已经引起科研范式的革命。科研模式变革是开展文献情报服务、科技情报服务的大前提, 决定了科研信息服务需求的形式和内涵。在这种科研模式变革中, 信息技术、智能技术的飞速发展与应用, 已经形成数据驱动的科研范式, 并且在不断深化和发展, 构成科研活动、知识发现产业化的新手段, 数字化的科研模式正在形成。未来的学术研究、知识发现, 以及对客观规律

的认知过程, 可能正在演变为信息计算或者数据计算的科研模式。《第四科研范式》详细描述了科学范式从经验范式、理论范式、计算模拟范式、数据发现(data exploration)范式的演变过程<sup>[1]</sup>。在数据发现的科研范式中, 将科学研究的理论、实验、模拟等统一, 构建一个整体化的科研发现平台, 科研数据通过仪器获取或者通过模拟的方式形成, 然后利用软件处理数据, 同时将信息/知识等存储在计算机系统中, 科学家利用计算平台分析和解析数据来完成科学研究过程。近年来, 数据发现式的科研范式已经成为客观现实, 如化学机器人、蛋白质折叠、AI在材料科学研究中的应用等, 都是以信息数据计算的方式进行科学研究。所以, 科研信息服务机构首先要把握这个趋势。

第二, 在新学术信息服务范式下图书馆、图情机构的定位与内涵也在发生一些变化。多年来图书馆界一直在适应环境和需求的变化, 不断创造和丰富服务内涵。例如, 在高校图书馆机构中, 积极倡导面向决策提供信息服务, 组织开展专项性的情报服务工作。国际图书馆机构同样呈现这样的发展态势。在图书馆学中, 很

多专家在倡导图书情报机构向智库研究和服务机构转型,也体现了图书馆服务内涵变化的一个特点。我们可以从这些变化当中仔细去分析、归结,提炼图书馆机构变化的特点和趋势。

目前,图书馆信息机构的工作对象在发生变化,体现了图书馆发展的时代特征。图书馆机构的工作对象简单来说面向图书、文献的,形成了面向文献的业务流程。信息机构进一步深化,以信息范式实现图书馆的管理与服务,将工作对象从文献上升到信息。归纳来说,现在图书馆和信息机构所面向的工作场景发生了变化,数据、信息、情报、知识、决策场景都成为工作对象的一些部分。如果按照信息分析这个视角,信息管理呈现为显性知识表达、隐性知识挖掘、情报表达以及报告形成,最后支撑决策服务。这体现了图书馆和信息服务机构的内涵定位的进一步变化。

除了在数据、信息、情报、知识和决策场景层面的内涵变化之外,图书馆/情报机构的工具方法也在发生演进,信息系统、内容管理、知识管理、情报分析、决策支持的技术工具不断得到深入应用。从图书馆、信息机构的工具演进技术上来讲,可以说信息化或者数据化导致了智能化应用。新一代信息技术的应用带来了变化,形成众多的工具、方法、平台等。同时,新工具与方法的形成,也造就了图情领域的理论范式的演化。笔者在《图书情报工作》2021年第1期撰写了一篇关于情报学范式变革的文章<sup>[2]</sup>,将其归结为信息论范式、信息管理范式、决策情报范式、信息咨询范式。这样的理论范式变化,可以说是信息管理从技术层面走向信息领域的内化过程,最后再走向社会化。所以,图情领域的理论基础和理论范围都是进一步扩大和推广。

在图情理论的各个范式下,其理论基础决定了技术路径,确定文献情报工作的服务形式,最后引领的发展趋势也各不相同。如在情报学的经典Brookes理论和知识方程指引下,形成了检索系统、知识地图,最后可以形成智能化的学术信息生态;在决策情报范式下,从决策情报理论出发,可以建立决策信息分析与服务、智库研究与服务的工作模式,形成智能化决策的信息支撑,进而演化为智能化决策或者数据决策服务的趋势;在社会化情报服务范式下,以商业信息服务为理论指导,在互联网、物联网、智能化网络支撑下,面向信息化社会的服务场景,发展了商业智能、社会化信息传播等服务,最后向社会智能大脑、智能信息生态等方面发展。

通过上面简单概要的分析,得出以下启示,对规划图书情报机构的未来具有借鉴意义。

第一,图书情报机构应该把数据和数据化建设作为工作的起点,树立数据思维,构建数据流程,发展数据管理的理论与方法,以数据为支撑,开展文献信息和情报服务工作。

第二,以数据和信息的数据化为原点,推动图书馆学情报学与数据科学的融合。这里要解释一下信息数据化的概念,按照一般的认识,学术研究中将数据变成信息,然后形成知识。但在文献情报工作中,我们搜集了大量信息,由于这些信息没有进行数据化,难以进行计算处理,也难以形成情报、形成支撑服务的新知识。因此,在文献情报工作中,信息的数据化应该是我们当前工作的难点,也是重点。

第三,当前文献情报工作应该以数据为起点,重构文献情报工作的价值链。在情报学领域,我们经常遵从数据-信息-知识-情报-智慧模型(DIKIW模型)<sup>[3]</sup>,构成一个完整的价值链。在信息管理领域,从信息这个中间体切入,使得文献情报机构的工作在服务链条或者信息价值链的人为断裂,在具体文献情报服务工作中,我们推动从信息端延伸到数据端,推动从信息端往知识、智慧这个方向延伸。如果文献情报工作内涵向智库扩展,需要将情报工作延伸到智慧情报(当然情报本身也有智慧、智能的含义)<sup>[4]</sup>。因此文献情报工作要以数据为起点重构服务价值链。

第四,全面推进文献情报工作的数据化,包括文献数字化、学术信息及实体信息数据化,社会信息、行为信息、管理信息、过程信息数据化。通过数据管理、数据融合、知识抽取、情报转换、决策支撑的理念来建立文献情报工作流程。

## 2 支撑“十四五”规划的基础和条件

中国科学院文献情报中心从1949年中国科学院图书馆成立起,不断向科研信息服务、知识服务方向演化发展<sup>[5]</sup>,现在进入了第三阶段,即建设国家科技知识服务中心。国家科技知识服务中心建设,起源于数字图书馆,发展面向科研的学科化服务,演化为科技创新知识服务,工作对象从印本信息资源建设到数字文献资源建设,发展到目前的科技信息大数据体系建设,形成支撑科技知识服务的数据基础。

中国科学院文献情报系统是由院级文献情报中心

和研究所所级的图书馆组成,是一个完整的工作体系。这个工作体系中有法人机构和非法人机构。其中,有3个法人机构,110多家研究所所级图书馆机构,还包括3家大学图书馆。采用全院协同的工作体系和工作模式,主要服务职能包括科技文献保障、科技创新信息支撑、战略情报服务、学科情报服务、产业科技情报支撑、学术信息交流、科学文化传播等。

中国科学院文献情报中心的服务对象包括中国科学院科研人员和全国科技工作者。在“十三五”期间拟定了3个“一线服务”,面向决策一线、科研一线和区域/产业发展一线。服务的是110多家研究所、3所大学,还有院党组这样一些科技决策部门。服务中重点强调三个方面,一是精准化服务,二是情报研究,三是精准化服务通过信息集成平台作为支撑。

## 2.1 形成了比较强的文献资源保障体系能力

在“十三五”期间,中国科学院文献情报中心重点进行文献资源的保障集成能力建设。一方面,形成和建立了安全可控的科技文献资源保障体系与服务能力。在资源保障体系中,从印本文献馆藏到数字化信息集成,再到开放学术信息的搜集,甚至延伸到新型的数据资源组织。在保障传统印本文献储备量持续增长的同时,形成一流的数字化信息保障能力。在开放信息资源组织以及预印本、机构知识库建设方面形成规模效应。如建立了OA全文库,通过对OA论文的搜集、整理、评价,形成大规模的全文库,搭建预印本平台,通过预印本平台搜集发表的学术论文。在中国科学院体系内建设了覆盖中国科学院所有研究所的机构知识库,搜集和存储全院科研人员所产生的所有论文、学术报告等。同时,进一步延伸和发展到社交媒体数据、遥感影像数据等。结合安全保障、安全服务能力的提升,我们也在探索如何与商业数据库出版商发展移动服务,来支撑这种科研和服务工作。在2020年疫情期间,在出版商的大力支持下,我们开通了移动服务,让科研人员在移动状态、在家里都可以完整获得科技文献。

另一方面,引领国家数字资源长期保存和论文开放获取的建设。在国家科技图书文献中心(National Science and Technology Library, NSTL)的大力支持下,组织、推动国家数字资源长期保存体系的建设。这个长期保存体系由NSTL的成员馆、高校图书馆(北京大学图书馆、清华大学图书馆等),以及军事科学院图

书馆等机构组成,甚至还包括中国图书进出口公司,都作为数字化文献信息的长期保存节点。目前正在筹建国家数字文献长期保存中心。同时,继续推动我国参与全球开放学术交流治理体系改革的建设,积极推动中国签署开放获取协议(OA协议),积极推动高能物理领域学术论文开放(SCOAP3),带领中国图书情报机构率先参加国际OA2020计划,来实现学术论文、学术信息的共享。

## 2.2 大力开展文献情报大数据平台及信息评价服务

“十三五”以来,中国科学院文献情报中心致力于建立科技文献情报的信息大数据平台,并开展数据型文献情报服务。第一,建成汇聚多源信息的科技文献和情报的大数据体系,这种多源信息可以把期刊论文、学术报告、学位论文等各类信息进行数据化处理,形成实体和知识关系,支持知识图谱、情报信息和情报知识的挖掘工作。第二,建立开发智慧型、计算型科技信息的服务平台,提供科技文献的移动获取、精准推送、智能挖掘和关联推荐等知识服务。同时,开发科技文献知识人工智能引擎,形成智慧知识服务平台。通过开发智慧科研平台、科技大数据知识发现平台等工作,形成数据采集、数据管理、自动监测、学者画像、机构画像和知识推送服务体系。第三,建设中文科技文献数据库(Chinese Science Citation Database, CSCD),积极开展学术信息评价服务。在大数据平台建设方面,CSCD建设是中国科学院的一项长期工作,已经形成数据库产品,在国际上销售和提供服务。目前CSCD已服务于200多家国内外高校,同时与国际平台合作,实现了CSCD平台服务国际化。在学术评价服务方面,依托文献信息开展学术论文、期刊质量的评价服务,包括OA论文的评价、论文撤稿的评价等。

## 2.3 形成了嵌入科研创新一线的学科情报与科研服务能力

第一,建立协同化的学科领域分工模式。在中国科学院体系内形成由院中心、成都中心、武汉中心、兰州中心、上海中心构成的情报服务体系,包括学科情报工作和产业技术情报服务等。全院文献情报体系按照领域、专业方向有一定分工,形成领域监测、领域态势把

握和情报分析等工作的方向。这是中国科学院自成立以来一直保留和坚持发展的科技情报服务体系。

第二,面向科技决策需要,组织战略情报研究与服务。在学科领域与方向发展态势监测的基础之上,开展战略情报扫描和形势研判,面向深化科技体制改革的科技战略问题开展研究,支撑情报决策。战略情报的服务对象十分广泛,通过专题报告、专业且深入的研究,支撑全院决策,以及支撑国家相关科技决策部门。

第三,对学科领域前沿与国际科技战略态势进行了系统的把握。系统整理国际科技领域的战略趋势前沿和创新政策发展态势,为决策部门提供分析报告和分析服务。

第四,重点发展了面向研究所科技创新和成果转化提供知识服务。其中,最大一项工作是支撑科研人员的科技信息获取,即科技创新的信息素质培训和教育。在科技信息素质培训和教育中形成了庞大的培训体系,如在中国科学院大学开设了40多门专业领域方向的信息素质培训课程,均为学分课程,可能是全球最大的信息素质教育培训体系。同时,结合学院的体系建设,将中国科学院大学的课程推向研究所、推向科研一线,将信息素质培训延伸到研究所课题组、实验室。因此,这是一个很庞大的信息素质培训体系。以深入科技创新场景的培训为基础,不断拓展服务,为全院100多家研究所提供学科情报服务,重点支撑研究所层面的学科战略规划、重点实验室评估以及成果转化和科技创新等服务工作。此外,还面向各个省科学院,面向区域发展行业提供服务。如依托于全国科学院联盟,建立了覆盖19个省科学院的信息与情报服务网络,依托于中国科学院的12个分院来开展与区域科技创新的合作,向区域和产业发展提供知识化服务。

第五,发展信息计算型知识服务,进一步实现信息深度挖掘与服务精准化。基于多维度数据计算和深度挖掘,形成对人才、高端科学家的立体刻画和认知,促进潜力科学家的发现以及科研团队的组织等。基于多源数据的信息计算模型来挖掘关键技术清单,如已经出版和组织了3本关于颠覆性技术创新的研究专著,即《颠覆性技术创新——信息技术领域》《颠覆性技术创新——生命科学领域》《颠覆性技术创新——能源领域》。其中第3本专著,基于论文数据、专利数据、社交平台数据等,形成能源技术领域的热点方向;然后结合专家的研判,共同形成能源领域颠覆性技术清单。上述情报研究工作,对于指导学科领域发展和领域规划

起到很重要的作用。当然,我们也开展一些关键技术清单的研究工作,如采用机器学习和文本挖掘的方式挖掘“卡脖子”技术清单等。

## 2.4 新型学术交流模式探索

“十三五”期间为探索科技期刊学术论文新出版模式,研究和发布国内首个《中国科学院学术期刊全文XML描述标签集》标准,目前在10多种期刊当中应用,这个标准是针对期刊学术论文的内容发布标准,希望形成期刊领域的一个全文论文发布的标准。同时,开展中国科技论文发布平台的建设,探索新型期刊的出版模式,探索了几种开放出版,建立数据出版工作,对于提高学术交流产生了很好的影响。

## 3 制订“十四五”规划的基本思路

基于对发展趋势的认识和判断,结合已有的发展基础,中国科学院文献情报中心在“十四五”期间确立阶段目标:面向“战略决策一线、科技创新一线、区域与产业一线”,以有效支撑科技创新高质量发展和科学战略决策为核心,强化中国科学院文献情报系统建设,建设国际一流的国家科技知识服务中心。在这种目标的体系下,来提升科技创新服务支撑、高端科技信息交流和决策战略情报支撑能力建设。具体任务包括强化大数据中心建设,形成自主安全、系统集成和关联融汇的智能化科技知识大数据平台;针对中国科学院科技创新的需求来深化学科服务的机制建设,建立服务科技创新的学科情报工作体系,来提高对科技自立自强的知识服务保障效率;建设支撑科技决策的情报感知与服务平台,形成数据、信息、情报、专家一体化的决策知识服务生态;集成数字学术信息评价、新型学术信息交流等能力,来建设中国科学院科研论文和科技信息高端交流平台。

从规划思路上讲,第一个方面,希望按照“强化核心、创新体制”的发展思路,系统升级数据、技术平台和情报等能力,强化对科技创新和战略决策的知识服务支撑。在工作上分成情报、文献、信息3个层面,按照数据化、平台化和精准化服务进行展开。这里需要特别强调的是,在数据基础上形成信息库、情报数据库,这是不断深化信息的数据化加工,将信息粒度进一步细化的过程。在平台方面更加强调智能化、智慧化,最后

是形成精准知识服务。这种精准实际上是一个智能化的信息保障和面向科技创新过程的、嵌入式科技创新过程的知识服务,还有支撑战略决策感知的一些情报服务工作。

规划思路的第二个方面,要结合中国科学院文献情报系统开展一些已有优势,发挥学科情报服务的优势和战略情报服务的优势,建立学科情报和战略情报的协同工作机制。在这种协同工作机制之下,可以形成决策信息的简报、专报或者信息数据化,建立一种集成研讨、趋势分析和态势感知的服务体系,支撑各类服务的要求。一方面,可以通过信息服务来支撑科技决策的需要,融合战略情报和学科情报研究,形成决策支撑的新服务机制。另一方面,在整个服务体系当中,整合战略情报力量,整合战略科学家、学科专家和数据管理者的能力,来提升情报服务的质量。

中国科学院文献情报中心的国家科技知识服务体系是围绕情报体系在数据、中台、集成智慧服务层面进行展开。基本任务包括建设科技文献与科技知识大数据中心,建设知识计算平台与工具体系,形成智能知识服务和智能情报研究架构。

在智慧服务数据建设方面,要建设大数据中心,即全面收集、保存和处理科技信息来支撑科技创新的科技文献大数据中心。除了需要对信息进行标注,不断拓展信息的数据加工精度,还要形成学科专业领域的分布式学术信息加工平台,形成整体的数据-信息-情报工作体系。对于信息和数据,要形成以学术知识图谱的方式对学术大数据进行管理,研发知识图谱等智能知识组织工具和平台,实现数据组织的智能化和流程化。

在智慧中台,加强数据计算和智能服务支撑能力建

设。可以建设新一代知识智能技术工具系统与平台,来提升服务能力,要充分利用以往的智能挖掘技术,甚至包括社会信息计算的能力,把充分计算信息能力引入文献服务体系平台建设当中。

要提供智慧情报服务,需建设智慧科情,即加强嵌入式的智能知识服务和智能情报研究。围绕科学研究和科研管理与决策的服务需求,提升科技文献、科技信息的服务水平。同时,把情报学科和学科情报服务嵌入全院科研管理项目和团队的科研过程当中,来开展数字化、数据化的智能知识服务。通过这样一种方式,建立嵌入学科主题的知识智能服务机制,提高知识服务的能力,把知识服务工具的应用延伸到数字化科研环境当中,来建设新型服务模式。

## 参考文献

- [1] HEY T, TANSLEY S, TOLLE K. The Fourth Paradigm: Data-Intensive Scientific Discovery [EB/OL]. [2021-03-03]. <https://www.doc88.com/p-777870574601.html?r=1>.
- [2] 刘细文. 情报学范式变革与数据驱动型情报工作发展趋势 [J]. 图书情报工作, 2021, 65 (1): 4-11.
- [3] AHSAN S, SHAH A. Data, Information, Knowledge, Wisdom: A Doubly Linked Chain? [C] // ICKE. Proceedings of the 2006 International Conference on Information & Knowledge Engineering. Las Vegas: CSREA Press, 2006: 270-278.
- [4] 曾建勋, 魏来. 大数据时代的情报学变革 [J]. 情报学报, 2015, 34 (1): 37-44.
- [5] 刘如, 吴晨生, 刘彦君, 等. 中国科技情报工作的传承与发展 [J]. 情报学报, 2019, 38 (1): 38-45.

## 作者简介

刘细文, 男, 1965年生, 博士, 研究员, 中国科学院文献情报中心主任, 研究方向: 情报学。

On the 14<sup>th</sup> Five-Year Plan of National Science Library of Chinese Academy of Sciences: Frameworks Design Based on the Notion of Data-Information-Knowledge-Intelligence

LIU XiWen

(National Science Library of Chinese Academy of Sciences, Beijing 100190, China)

Abstract: The article analyses systematically the challenges and trends of library and information science institutions in the new ages and objects, connotations, theories, technologies, and explains what are the impacts and references on the LIS services by the digital research model and changes of service need. In the design of future LIS service, we should replace the information with the data as the start point of LIS workflows. Then the article concludes roundly the development of National Science Library of Chinese Academy of Sciences in recent years, such as the support capacities of academic literature and information, capacities of platform management and data mesh-up, service capacities for researches and S&T development, academic communications and information assessment. In the period of 14<sup>th</sup> Five-Year Plan, National Science Library of Chinese Academy of Sciences would provide services in the "strategical decision line, S&T innovation line, region and industrial innovation line", develop the strategy intelligence and technical intelligence services, and then establish the center of national S&T knowledge services.

Keywords: National Science Library of Chinese Academy of Sciences; 14<sup>th</sup> Five-Year; Development Plan

(收稿日期: 2021-05-10)