

# 科研学术大数据的精准服务架构设计\*

谢靖<sup>1,2</sup> 钱力<sup>1,2</sup> 师洪波<sup>1</sup> 孔贝贝<sup>1</sup> 胡吉颖<sup>1</sup>

<sup>1</sup>(中国科学院文献情报中心 北京 100190)

<sup>2</sup>(中国科学院大学图书情报与档案管理系 北京 100190)

**摘要:**【目的】研究面向科研学术大数据的精准服务架构,解决科研学者获取精准知识服务的问题。【方法】通过分析互联网应用精准服务现状,从数据组织、技术方法、应用场景三个方面,归纳对比精准服务的方法,面向科研学术生态链条设计精准服务体系。【结果】精准服务架构设计梳理关联了数据生产、技术研发、应用服务环节,支撑了科研学术的精准搜索与推荐服务应用。【局限】该架构未与传统方法对比,还需结合实际应用评测验证精准性提升效果。【结论】本文解决了精准服务技术方法的单一性问题,为应用系统研发提供系统的架构支撑。

**关键词:** 精准服务 学术大数据 架构设计 用户画像

**分类号:** TP391

**DOI:** 10.11925/infotech.2096-3467.2018.1366

## 1 引言

科研学术数据爆炸式增长,以文献检索为核心的传统学术服务模式不能满足用户对精准知识服务的诉求。中国科学院在“十三五”规划中将精准科研学术服务列为主要突破项目,设计精准服务架构:囊括数据组织、知识图谱、语义搜索、用户画像、主动推荐等前沿技术,将各类问题难点技术方法串联,形成一套面向科研学术应用服务的精准服务体系。

本文结合科研人员的需求,分析对比现有精准服务应用系统;贯通数据组织、技术方法、应用模式三个重要环节要素,设计科研学术大数据精准服务架构,支撑科研学术大数据精准服务生态链条。

## 2 现状分析

### 2.1 互联网平台精准服务环境

#### (1) 基于召回率的用户兴趣跟踪

Netflix 视频构建标签系统<sup>[1]</sup>为一个视频内容打不同标签,并对标签设定权重。根据向用户推荐的内容,观察用户点击浏览行为,通过对内容标签的召回,发

现用户的兴趣,从而构建闭环迭代的用户兴趣跟踪系统。

#### (2) 用户画像与内容画像结合

今日头条构建了用户画像和内容画像两套画像系统,通过机器+人工结合的方法给每个用户进行分类和打标签<sup>[2]</sup>,对每条资讯分类和打标签。基于两套画像系统,将分类一致、标签匹配度最高的内容推荐给用户,从而实现千人千面的个性化资讯推荐服务。

#### (3) 群体行为特征分析的内容推荐

淘宝、京东等国内热门的购物网站时刻对用户群体的购买行为进行收集分析,通过构建用户与购买物品的大数据关系网络,从而发现用户购物兴趣特征。常用的如“协同过滤算法”发现相似的用户,“基于物品特征的过滤算法”发现相似的物品<sup>[3]</sup>。当前已经构建起一套成熟的推荐系统为用户推荐符合购买兴趣的物品。

### 2.2 科研学术界精准服务环境

#### (1) 基于科研社交关系的精准服务

国外的 ResearchGate 和 LinkedIn 最早开展科研圈社交关系服务,通过学术合作和学术社交关系将科研

通讯作者: 谢靖, ORCID: 0000-0001-6698-1786, E-mail: xiej@mail.las.ac.cn。

\*本文系中国科学院文献情报能力建设专项项目“基于大数据计算的知识发现服务平台建设”(项目编号: 院 1853)的研究成果之一。

人群分类,将科研圈合适的人及科研活动以社交网络的方式精准推送给合适的学者。近期,国内创业公司研发的科学家在线(ScientistIn)产品,以科研学者为中心,面向科研成果转化,构建学者和产业化公司的社交合作网络,目标是建立精准的“学者-产业”的科研成果转化的桥梁。

### (2) 基于文献成果管理的精准服务

通过学者公开发表文献,提取用户研究兴趣,提供精准服务是科研学术界的另一特点。Mendeley 和微软学术都是基于用户文献的服务平台,挖掘用户兴趣方向并从多个维度提供用户精准服务。国内百度学术和中国知网注重中国学者及中文文献的管理和挖掘,研发学者主页,收集发现用户学术动态,从而提供附加的精准文献推荐服务。

### 2.3 精准服务应用对比分析

互联网公司的精准推荐注重数据内容层面计算,目标明晰,方法单一。但精准度、环境依赖性带来的推荐陷阱问题亟待解决。三种不同算法各有优缺点,需要多种维度和方法有效结合使用。而学术界精准服务科研社交关系利用学术界同行评价的优势<sup>[4]</sup>,使优

势资源得以发掘。但是社交网络的发展有更多随意性和自发性,精准程度受到社交关系的限制。文献成果管理的精准服务以学者为核心,注重人的数据挖掘计算,文献数据更具有权威性,从而使得精准程度稳定可靠。但是文献发表具有周期性和滞后性,对实时性精准服务不利。

据此,精准服务需要设计一种统筹的系统架构,以科研学术应用场景模式为导向,将优势的数据组织手段、计算机技术进行整体规划,以更好地驱动科研工作者的知识化、智能化的精准服务。

## 3 架构设计

科研学术大数据精准服务主要面向科研学者与科技情报人员,提供精准的数据服务与应用服务支撑。精准服务架构设计必须以应用服务为核心目标,数据组织与技术方法是两大主要支撑与驱动力。在精准服务架构中,应用服务的需求促进数据组织的进化,促进技术方法的演进,从而推动精准应用服务的升级。

如图 1 所示,本文设计的精准服务体系架构分为三个层次:应用服务层、技术方法层和数据组织层。

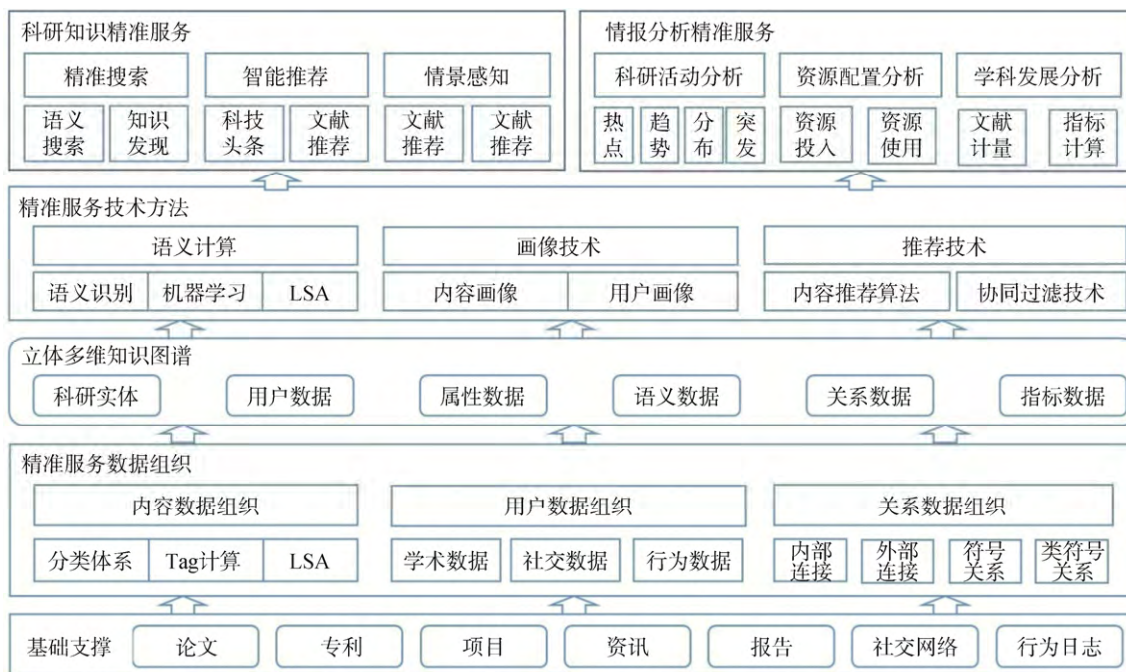


图 1 科研学术精准服务架构

### (1) 应用服务层

用户需求是精准服务架构设计的核心导向。根据

需求调研,将用户需求分为科研用户的知识服务需求和管理决策的情报分析需求。面向科研用户,精准应

用服务需求从文献服务转变为语义精准搜索、智能知识推荐和根据用户个人当前环境的情景感知服务；面向情报分析与决策者需要掌握科研活动、资源配置、学科发展等指标计量分析的诉求，精准服务从人力密集型的数据收集、统计、分析转变为数据密集型的智能精准计算服务支持。

#### (2) 技术方法层

语义计算技术通过深入挖掘知识层面的实体与关联关系，为精准服务组织一系列相关的知识原料；画像技术面向科研用户个人，精准分析行为与兴趣特征，精准定位要在何时、何地、提供哪方面的知识；推荐技术构建知识到用户的关系桥梁，转变“人找知识”为“知识主动找人”。

#### (3) 数据组织层

数据组织层计算加工原始基础数据，最终形成可支持精准服务的立体多维知识图谱。数据组织包含内容数据、用户数据与关系数据的组织计算。内容数据组织提供精准服务的知识素材；用户数据组织支撑深度分析用户需求特征；关系数据组织架构知识与用户的桥梁。

## 4 数据组织

数据的精准组织是利用数据分类、标签计算、主题标引等方法将数据转变为有序有规则的组织结构，使其具备提供精准服务的基础支撑。面向学术的精准服务包括：内容层面数据组织，用户层面数据组织，以及内容和用户之间的连接关系三个重要环节。数据组织的精细度决定了服务的精准程度。

### 4.1 内容数据组织

#### (1) 内容数据组织原理

分类体系：全领域学科的分类体系，如中图分类法等。采用树形分类层级结构，层数越深，粒度越精准。而这种分类法并不能完全适用于专业领域，如微生物学专业领域的分类结构<sup>[5]</sup>。分类体系中需要两者的结合应用。

标签系统：Web2.0 时代诞生的文本标签系统，不再局限于传统树形分类体系。灵活的长短文标签可由用户自发产生，也可由计算机算法快速生成，迎合互联网大数据时代热词迭代变化的节奏。

语义组织：拆分文本语义内容，从认知层面上识

别与挖掘数据。如 NCBI 的 MeSH 利用领域词表深度挖掘提取医学领域知识，KOS<sup>[6]</sup>利用知识组织体系对文献信息深度标引。突破分类和标签表层组织文献数据的局限，实现文本内容的深度精准挖掘。

#### (2) 内容数据组织架构方法

内容数据组织应采用递进式组织架构：首先内容必须从属于一种分类，确定该内容的基本层级与位置。用于推荐系统启动时，进行基本的内容粗粒度的推荐；其次，建立内容自动标签系统，以支持内容的多维度分类，打破内容单一分类的局限性，增强推荐应用的多样性和精准性；最后以文本语义拆分，实现具体知识单元的提取，实现内容在段落、句子以及具体知识对象的定位，以增强细粒度的语义认知层面上的精准内容呈现。

递进式数据精准组织方式在中国科学院文献情报中心“十三五”规划项目“慧科研”精准推荐应用示范系统中得以试验验证，采用递进式内容组织架构可以更多地避免内容噪音以及语义消歧带来的内容二义性，降低一定的技术和算法的复杂性。

### 4.2 用户数据组织

基于科研学术人员调研，将人员的特征信息分为 4 个大类：基本属性、行为日志、发表文献和社交活动，如图 2 所示。4 种类型数据随时间而动态变化。因此数据收集既要概括全貌，更要跟踪当前活跃信息。

#### (1) 科研用户数据特征分析

基本属性信息：涵盖年龄、学历、职称、教育经历、工作经验等科研学术用户的基本信息，代表个人的学术成长、科研能力、兴趣方向，是精准服务精准定位。需要设定数据渠道和使用权益，保护个人隐私，突破数据获取计算的障碍。

行为日志信息：通过收集用户的上网浏览内容等行为，对日志数据进行挖掘分析，发现用户兴趣与潜在需求。通常包括：上网环境、行为动作、关注主题、时间分布与地理分布等。行为日志主要特征是数据碎片化，需大量数据关联计算才能达到精准服务的要求。

发表文献信息：收集整理用户公开发表的文献信息，通过文献计量准确计算掌握用户的兴趣特征与学术水平。文献信息计算精准性高，弊端在于文献发表的滞后性，不能及时发现科研用户前沿动态，影响精准服务的时效。

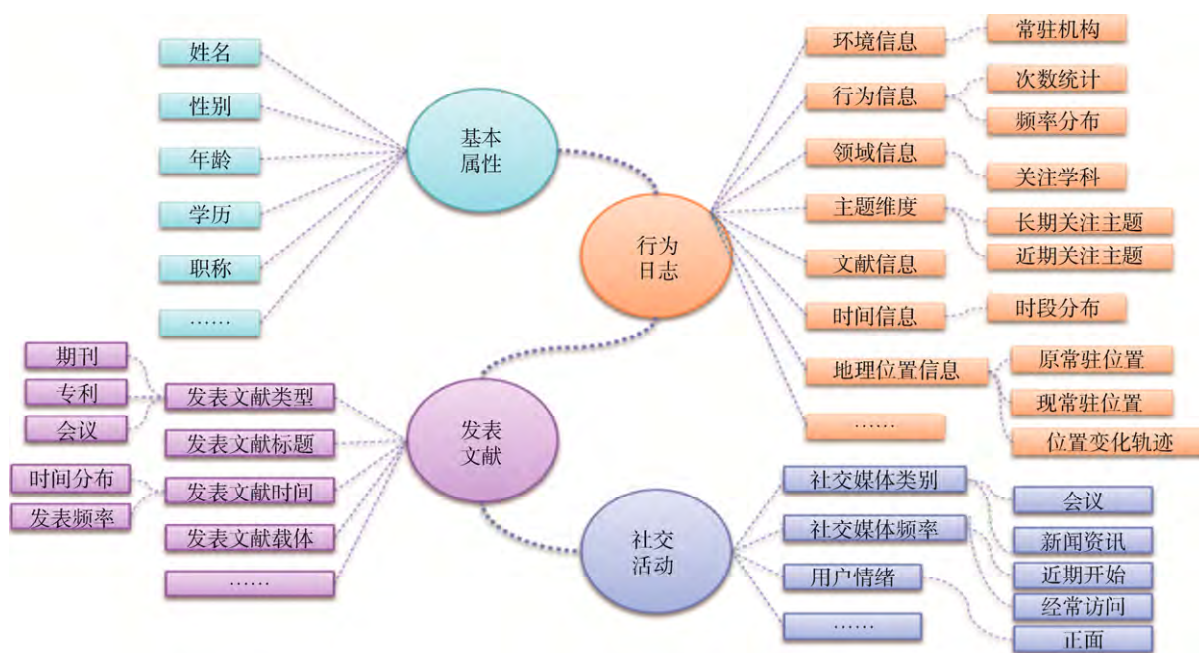


图2 用户数据精准组织结构

社交活动信息：科研学术用户在社交媒体发表的言论、转发、评论等交流传播活动。这些数据信息可以有效克服文献发表的滞后性，是精准服务重要的数据补充。

#### (2) 科研用户数据组织架构方法

用户数据组织架构见图2，以左侧用户基本属性与发表文献信息为基础，勾画用户基本画像标签数据，为用户推荐提供长期的兴趣方向计算依据；以右侧行为日志数据与社交活动计算当前行为特征，并通过动态反馈的方式，影响用户兴趣标签的计算。用户长期兴趣数据计算与短期行为数据计算结合的组织方式，既能够及时把握用户当前兴趣活动，又能够指引基本兴趣方向，克服行为日志不断强化当前行为导致的推荐计算陷阱。

#### 4.3 关系数据组织

科研学术关系数据包含内部连接、外部连接。内部连接在内容层面上体现为内容数据之间的知识关系；在用户层面上体现为用户之间的设计关系，如项目合作、论文共著、相互关注等。外部连接指内容与用户之间的关联，如用户发表文献、用户承担项目等；而经过分类聚类、Word2Vec 等类符号化计算得到的内容相似性等隐性连接，并不能显性表述其连接意义，但是在精准服务计算中依然承担着重要的角色。科研

学术大数据精准服务必须组织应用好关系数据，特别是机器学习计算得到的隐性连接关系数据。

关系数据起到桥梁的作用，将数据知识化组织为图谱。科研学术知识图谱中的关联关系是支撑精准应用服务的重要基石。在应用架构中应以显性的确定关系为主，用已知的确定关系作为首选依据；以类符号化计算的隐性关系为辅助，补充缺失或者潜在发现的知识连接，丰富图谱关系。

### 5 技术方法

#### 5.1 用户画像技术

##### (1) 技术思想

交互设计之父 Alan Cooper 最早提出用户画像 (User Profile) 的概念<sup>[7]</sup>，旨在勾画目标用户的诉求，从而有效设计构造产品原型以满足个性化需求。用户画像技术以人为本，深挖个性需求特征，从而为精准服务奠定基础。用户画像在精准司法、精准医疗、精准营销等服务方面都有广泛应用。面向科研用户学术画像，目标是实现对科研学术用户个人深挖用户行为特征，分析构建用户兴趣模型。

##### (2) 实现流程

用户数据收集：从时间分布、地理位置、网络环境、操作行为以及在网络行为中关注的领域、主题、

学者、出版物等多个维度刻画用户学术画像模型。

**用户特征挖掘:** 基于用户数据、利用计量统计学方法, 提取共性与特性的用户特征(如某一主题内容的统计浏览次数等), 建立用户特征描述素材。

**画像模型构建:** 深入分析科研用户特征素材, 结合相应领域知识库, 从用户特征数据中理解用户的习惯和兴趣, 从特征数据中统计归纳形成用户画像模型。

**精准定量计算:** 用户画像传统意义上可理解为给用户打标签, 然而用户的特征标签庞杂, 仅使用标签无法满足精准计算, 需要利用概率模型的代表方法支持定量的精准应用服务。

## 5.2 语义计算技术

### (1) 技术思想

语义计算技术可以让机器认知文本内容, 以支撑精准知识搜索服务。主要包括语义识别技术和语义索引技术。语义识别技术实现文本中语义知识对象的标注与知识关系的提取; 语义索引技术通过语义知识对象和语义关系创建语义索引, 支撑语义层面的精准搜索应用。

### (2) 实现流程

**文本语义挖掘:** 实现对文本内容中的知识对象标注与认知, 利用词表、语义本体等虽然能够实现半自动标注, 然而需要大量人工维护更新词表本身, 不适应大数据发展要求。利用 LSTM 机器学习在少量人工语料的训练基础上实现完全的自动标注计算, 在领域内被广泛应用。

**构建语义索引:** Scott Deerwester 等提出 Latent Semantic Analysis(LSA), 创建了“矢量语义空间”提取文档的“知识概念”, 分析文档与词之间的语义关系<sup>[8]</sup>。LSA 开创语义索引应用模式, 实现在语义空间的检索模式。语义索引与传统索引结合使用实现语义内容与文献字段的按需查询。

利用 LSA 创建的语义空间, 如图 3 所示。语义检索应用需要返回精准的兴趣相似的文章。0 视为用户本身兴趣位置, q 为用户检索关键词, q 周边分布为矢量语义空间, 待检索文章为 d1-d5。语义检索路径即从 0 出发, 检索命中为历史词汇与 0 连接线覆盖的三角形区域。检索结果排序标准为 d1-d5 到 0-q 直线的垂直距离, 如 d5,d1,d2。

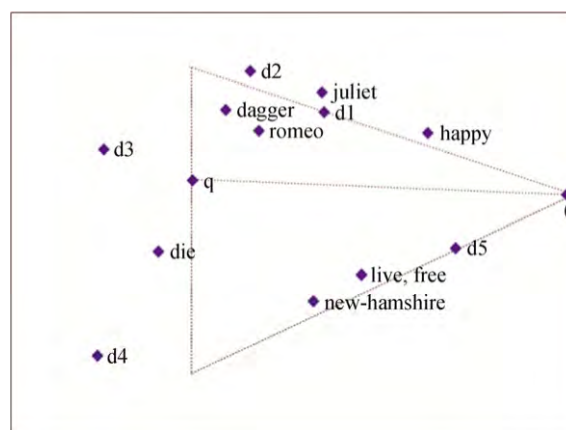


图 3 潜在语义分析精准检索示意图<sup>[9]</sup>

## 5.3 精准推荐技术

### (1) 技术思想

精准推荐技术连接内容与用户, 实现内容精准主动找人: 推荐技术诞生了基于关注热度推荐、协同过滤推荐、基于内容推荐、机器学习模型推荐和混合推荐等较多的算法<sup>[10]</sup>。各种算法具有其独特的优点和应用方式, 其中应用较多的是基于内容推荐与协同过滤推荐。

### (2) 实现流程

**推荐启动:** 用户进入应用系统之初在画像数据的基础上, 使用基于内容推荐算法, 通过标签及权重匹配, 将内容与兴趣相似的用户建立关联。可有效打破推荐系统冷启动问题, 实现了精准的信息递送。适用于频繁更新的论文、资讯等数据信息推荐。

**协同过滤(Collaborative Filtering)计算:** 当用户在系统中留下一定的日志行为记录, 协同过滤算法辅助寻找与用户自身兴趣、行为最相近的用户群体, 把他们关注的内容推荐给用户。该算法最大概率地推荐用户感兴趣的内容, 算法适用性较强, 适合于期刊、项目、学者等不频繁更新的内容推荐。

**权重计算:** 热度推荐、协同过滤推荐、基于内容推荐等多种算法需要混合计算, 得出最优的计算指标。权重计算需要根据用户应用需求, 面向不同的推荐资源进行不同的加权算法。根据具体需求和测试反馈, 不断调整权重参数, 让更适用的资源权重更高。

**智能排序:** 在应用阶段需要根据各种不同资源的混合计算, 智能计算出不同类型资源推荐的 Rank 值, 从而将满足用户需要的推荐信息排在前列。

## 6 应用服务

### 6.1 面向知识服务的应用

#### (1) 精准搜索

科研学术知识图谱可以提供对科研实体、知识概念与关联关系的数据支撑。精准检索系统可以根据用户画像,理解检索目的,自动精准导航筛选信息类型、

学科分类、主题标签。图 4 为中国科学院文献情报中心“十三五”重点突破项目研发的“科技大数据知识发现平台”,基于数据精准组织与知识图谱,实现了在论文、资讯、专利、项目、学者等多维度精准搜索的应用示范。随着语义索引研究的深入,从语义理解的方向为用户精准发现其需要的科技知识内容,逐渐由研究走向实际应用。

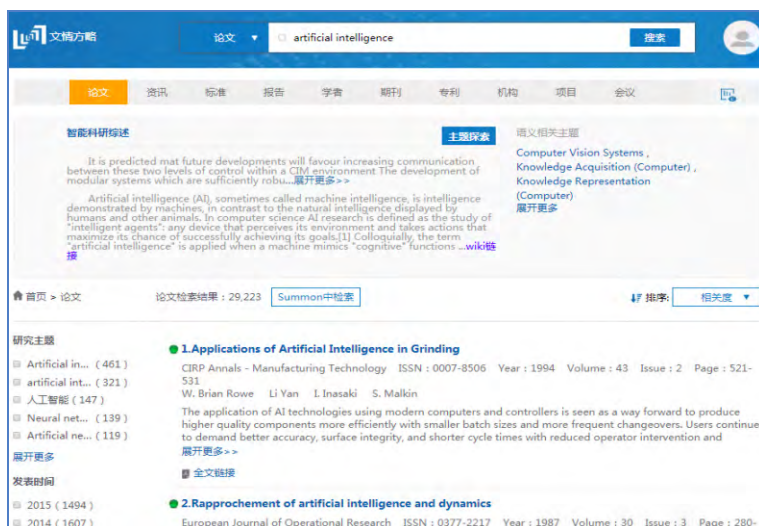


图 4 精准搜索应用系统应用示意图

#### (2) 智能推荐

推荐系统实现了知识主动找人的服务,逐步在各类型精准服务上得以应用,如推荐期刊、推荐学者专家、推荐文章专利、推荐学术会议等。特别是移动互联网时代,推荐系统在微信等智能 APP 的新信息渠道

中,成为科研学术用户发现科技前沿、跟踪领域研究进展、了解业界动态必不可少的工具。图 5 为中国科学院文献情报中心研发的“慧科研”精准服务示范系统。基于科研用户画像自动计算用户兴趣导向,实现对个人主动精准知识推送服务。



图 5 精准推荐应用系统应用示意图

(3) 情景感知

智能机已成为人们的随身助理, 强调实时性与场景化是科研学术精准服务的要求。在智能手机的协助下, 可以根据用户的行为习惯、地理位置感知、日程安排等环境信息, 在用户需要的时候给与精准的信息服务使得精准服务更加智能方便。例如感知用户的作息, 在工作的时间推荐科研文章, 在休假的时刻推荐感兴趣的休闲书籍; 计算科研用户出行目的地差旅标准, 提示用户科研项目申报、参加会议活动等情景感知智能服务。

6.2 面向情报分析的应用

(1) 科研活动分析

科技决策管理人员需要精准地把握科研群体动向, 支撑精准决策的判断。基于科研人员画像数据组织, 建立面向用户群体的精准分析系统平台。可以从投入、产出、影响力等多维指标分析, 掌控全体科研人员当前的人才梯队、研究热点、科研投入、发展趋势; 可以从全体科研用户、指定学科领域、指定国家地区、指定机构等多个视角感知当前群体科技人员行为动态。如图 6 所示, 国家科技图书文献中心(NSTL)研发的“科研用户学术画像大数据分析系统”, 实现了对 NSTL 用户日志行为数据的多角度全景监测分析, 能够支撑用户服务质量评估与领导决策。



图 6 精准搜索应用系统应用示意图

(2) 资源配置分析

科研学术大数据可实现从采购经费、人员管理等方面了解资源的投入情况。但是传统资源配置管理方法无法与资源利用情况、成果产出建立关联, 导致资源配置不能支撑应用与产出最大化。因此精准服务需要用户数据与内容数据的关联计算, 实现资源使用情况全方位统计, 让管理决策者实时掌握资源利用动态, 使资源采购优化前置, 最大程度上以应用需求为准绳实现精准的资源优化配置。

(3) 学科发展分析

学科情报分析人员当前主要的工作时间投入到

对数据的收集、整理与指标计算。在科研大数据环境下, 需要以数据密集型的计算环境支撑情报分析的数据收集、关联分析、精准指标计算等定量的计量工作。以数据为基础支撑, 解决学科发展分析中的数据完整性、准确性、计算维度的多样性等问题, 达到精准服务情报分析的目的。图 7 为中国科学院文献情报中心为某研究所研发的“机构学术分析系统”, 支持机构科研投入、成果产出、人才培养等方面数据精准分析, 更大程度上缩减情报人员的数据处理计算投入, 其将更多的时间投入到智慧密集型的情报分析工作中。



图 7 精准推荐应用系统应用示意图

## 7 结 语

本文旨在解决传统精准服务片面关注数据或者技术方面带来的局限性的问题，在总结分析当前互联网产品及学术服务应用的基础上，设计了科研学术大数据的精准知识服务架构。从数据、技术、应用三个层面详细阐述精准服务的逻辑与方法，构建精准服务分层体系架构，为应用系统研发提供理论与架构支撑。而本文设计的精准服务架构涉及内容不够完善，未来需要与传统方法对比分析，基于该架构提供精准服务的效用还需结合具体应用进一步评测验证。

### 参考文献:

[1] Gomez-Uribe C A, Hunt N. The Netflix Recommender System: Algorithms, Business Value and Innovation[J]. ACM Transactions on Management Information Systems, 2016, 6(4): 1-19.

[2] 曹欢欢. 今日头条算法原理[EB/OL]. [2018-01-16]. <https://baijiahao.baidu.com/s?id=1589714822818540008&wfr=spider&for=pc>. (Cao Huanhuan. Algorithm Principle of Today's Headline[EB/OL]. [2018-01-16]. <https://baijiahao.baidu.com/s?id=1589714822818540008&wfr=spider&for=pc>.)

[3] 洪亮, 任秋圆, 梁树贤. 国内电子商务网站推荐系统信息服务质量比较研究——以淘宝、京东、亚马逊为例[J]. 图书情报工作, 2016, 60(23): 97-110. (Hong Liang, Ren Qiuyuan, Liang Shuxian. A Comparative Study of Information Service Quality of E-commerce Sites' Recommender Systems ——

Take Taobao, Jingdong and Amazon as Examples[J]. Library and Information Service, 2016, 60(23): 97-110.)

[4] 王贤慧, 袁军鹏. 一种面向社会关系的同行评议方法[J]. 科技管理研究, 2017(23): 228-232. (Wang Xianhui, Yuan Junpeng. A Study on Peer Review by Experts Based on Social Relationship[J]. Science and Technology Management Research, 2017(23): 228-232.)

[5] York A. Next-generation Bacterial Taxonomy [J]. Nature Reviews Microbiology, 2018, 16(10): 583.

[6] Tudhope D, Nielsen M L. Introduction to Knowledge Organization Systems and Services [J]. New Review of Hypermedia and Multimedia, 2006, 12(1): 3-9.

[7] 陈慧香, 邵波. 国外图书馆领域用户画像的研究现状及启示 [J]. 图书馆学研究, 2017(20): 16-20. (Chen Huixiang, Shao Bo. The Research Status and Enlightenment of the User Profile in the Library Field at Abroad [J]. Researches in Library Science, 2017(20): 16-20.)

[8] Dumais S T. Latent Semantic Analysis [J]. Information Science and Technology, 2004, 38(1): 188-230.

[9] Thomo A. Lantent Semantic Analysis (Tutorial) [EB/OL]. [2013-10-01]. <http://openscienceasap.org/wp-content/uploads/2013/10/Latent-Semantic-Analysis-Tutorial.pdf>.

[10] Francesco R, Lior R, Bracha S, et al. Recommender Systems Handbook[M]. Springer US, 2015: 1-34.

### 作者贡献声明:

谢靖: 设计精准服务架构, 撰写、修订论文;



钱力: 提出精准服务架构研究思路, 论文修改;  
 师洪波: 提供精准服务技术方法研究内容;  
 孔贝贝: 提供精准服务数据组织研究内容;  
 胡吉颖: 提供精准服务应用模式研究内容。

### 利益冲突声明:

所有作者声明不存在利益冲突关系。

收稿日期: 2018-12-04  
 收修改稿日期: 2019-01-04

## Designing Framework for Precise Service of Scholarly Big Data

Xie Jing<sup>1,2</sup> Qian Li<sup>1,2</sup> Shi Hongbo<sup>1</sup> Kong Beibei<sup>1</sup> Hu Jiying<sup>1</sup>

<sup>1</sup>(National Science Library, Chinese Academy of Sciences, Beijing 100190, China)

<sup>2</sup>(Department of Library, Information and Archives Management, University of Chinese Academy of Sciences, Beijing 100190, China)

**Abstract:** [Objective] This paper proposes a framework for precise service of scholarly big data, aiming to improve knowledge acquisition of researchers. [Methods] First, we analyzed the status quo of online precision services. Then we summarized and compared the methods of precision services from the perspectives of data organization, technical methods and application scenarios. Finally, we designed the framework for academic eco-chain of scientific research. [Results] The framework connected data production, technology research and application development, which supported the precise search and recommendation of sci-tech data. [Limitations] More research is needed to evaluate the framework with real-world cases. [Conclusions] This proposed framework could help us build better academic precision search systems.

**Keywords:** Precise Service Scholarly Big Data Architecture Design User Profile

### 2019 人工智能发展趋势预测

VentureBeat 网站采访了多位业界杰出人士, 他们通过多年与全球最大的科技和工业公司合作, 了解正确的人工智能方式。

(1) Google Brain 联合创始人 Andrew Ng

①2019 年, 人工智能将被用于科技或软件公司以外的地方。下一波大规模的价值创造将会是制造厂或者农机设备公司、医疗保健公司开发各种 AI 解决方案来帮助自身企业的时候。

②2019 年, 预计人工智能 / 机器学习领域会取得有助于整个领域发展的进展, 包括 AI 可以使用更少的数据得到更精确的结论, 也就是所谓的“小样本学习”。

(2) Facebook AI Research 创始人 Yann LeCun

2019 年, 人工智能系统将在灵活性上取得进展, 并且能够在事件之间建立因果关系, 也就是说不仅是通过观察来学习的能力, 而且是通过实践来理解的能力。

(3) 埃森哲人工智能全球负责人 Rumman Chowdhury 博士

①2019 年, 政府会对全球技术界有更多的审查和监管。AI 以及全球科技巨头所掌握的权力, 引发了诸如该如何监管该行业和技术的诸多问题。2019 年, 将不得不开始寻找这些问题的答案——当一项技术是一种具有特定背景结果的多用途工具时, 将如何对其进行监管? 如何建立监管体制才能不妨碍创新或偏袒大公司(谁能承担合规性成本)而不是小初创企业? 我们应该监管到什么层面? 国际性、国家性还是本地性?

②2019 年, 人工智能在地缘政治事务中的角色还将继续演进。这不仅仅是一项技术, 更是经济和社会的塑造者。人工智能之争不仅仅是计算机能力和技术智慧的较量, 就像冷战不仅仅是核能力的较量。我们有责任以一种更公正、更公平的方式去重新创造世界, 现在我们有这种难得的机会。

③2019 年, 人工智能将在家庭中得到更多应用。许多人已经变得更加习惯于使用像 Google Home 和 Amazon Echo 这样的智能音箱, 以及很多智能设备。

(编译自: <https://venturebeat.com/2019/01/02/ai-predictions-for-2019-from-yann-lecun-hilary-mason-andrew-ng-and-rumman-chowdhury/>)